Volume 2 *March, 1932* *Number 3*

PHYSICS

SOME ASPECTS OF ELECTRICAL PROSPECTING APPLIED IN LOCATING OIL STRUCTURES

By Leo J. Peters and John Bardeen
The Gulf Companies, Research Department, Pittsburgh, Pa.

(Received February 2, 1932)

Abstract

Electrical prospecting is defined as the science and the art of determining the variations of the electrical constants (resistivity, magnetic permeability and the dielectric constant) of the earth's crust and of interpreting these variations in terms of geological structure. The most successful systems of prospecting are based on the study of resistivity variations. The basic assumption made is that in general changes of resistivity follow the bedding planes. Electrical methods of exploration may be divided into two classes, direct current methods and alternating current methods. In part II the fundamental theory of direct current method is discussed and a typical survey is described. Part III deals with the theory of alternating current methods, with particular reference to the optimum frequency to be used. It is shown that in general very low frequencies are desirable. Two alternating current surveys made by the Swedish American Prospecting Corporation are briefly described. In the conclusion (part IV) some of the difficulties of electrical prospecting are discussed. The depth to which investigations may be carried is limited. In the present stage of the art, it would take exceptionally favorable conditions to obtain reliable information much in excess of 2,000 feet. However, it is stated that improvements in methods of interpretation and in field technique should give electrical methods a definite field of usefulness in prospecting for oil.

I. Introduction

ELECTRICAL prospecting may be defined as the science and the art of determining the variations of the electrical constants of the earth's crust and of interpreting these variations in terms of geological structure or in terms of mineralization of a region. The three electrical constants whose variation may be determined are the resistivity, the magnetic permeability and the dielectric constant.

The resistivities of earth materials vary from a fraction of an ohm·cm to 10^{17} ohms·cm. The materials having values below 100 ohms·cm are confined largely to mineral ores such as marcasite, galena, molybdenite, chalcopyrite and iron pyrite. In the study of oil structures the resistivities encountered vary from 100 to 10^{17} ohms·cm. In the Texas Gulf Coast region the common variation is from 200 to 20,000 ohms·cm with the exception of rock salt of which the resistivity has a value of the order of 10^{17}.

The dielectric constant of earth materials has a value which ranges from 1 to 10. The variation in magnetic permeability, except in rare cases, is negligible.

Systems of electrical prospecting are, therefore, confined to studying the manner in which the resistivity or the dielectric constant varies in the earth's crust. The relative effect which these two constants have on the measurable electromagnetic quantities at the surface of the earth will be discussed later. However, it can be stated that up to the present time, the most successful systems of prospecting have been based upon a study of the resistivity variations in the crust of the earth.

The problems of any scheme of geophysical prospecting may be stated as follows:

First, from mathematical and physical considerations it is decided what data must be taken at the surface of the earth in order to predict the distributions of a chosen physical constant or constants in the earth's crust.

Second, instruments and field technique are developed to obtain these data.

Third, mathematical tools are developed which translate the quantities measured at the surface of the earth into the distribution in the earth's crust of the physical constant or constants under consideration. When these physical and mathematical problems are solved there remains the problem of finding out what relation (if any) the distributions of the physical constants which are determined bear to geology and of interpreting the measurements in terms of geology. Furthermore, to be economically justifiable the overall cost from field measurements to translation into geology must be commensurate with the value of the information obtained.

An ideal system of prospecting for oil would be one which directly indicated the presence of oil. Many attempts have been made to devise an electrical system which would meet this ideal requirement. However, these attempts have met with little or no success. In the present state of the art, the practical systems of prospecting for petroleum are those which are designed to give information concerning subsurface geology.

The basic assumption in mapping geological structures by electrical methods is that in general changes of resistivity follow the bedding planes. Thus a clay bed or shale bed may have a lower resistivity than surrounding limestones or sandstones. These beds usually are parallel with or are only slightly inclined to the surface of the earth. In general, the lateral changes in resistivity are neglected, and the changes in resistivity with depth are determined. If many such determinations are made over an area, the resistivity depth curves may be correlated and the subsurface topography may be obtained.

Generally one or more beds which have a large resistivity contrast with the beds above or below them are used as marker beds in mapping geological structures. It is obvious that the bed which is being used as a marker must conform to the bed carrying the oil in order to be of any practical use as a marker bed in exploring for petroleum. In many localities the relatively shallow beds are more or less parallel to the deeper ones. However, in many other places major unconformities occur and the shallow beds are no longer substantially parallel to the deeper ones. In such cases it is necessary to follow a

marker bed which is below the unconformity and is conformable to the oil bearing horizon. The depth to which it is possible to work with any geophysical scheme is, therefore, an important consideration.

The depth which can be obtained reliably in electrical prospecting is a much debated question. Up to the present time the depth at which structures have been reliably mapped by electrical surveys has not exceeded 1500 feet. Electrical surveying has gone but a little deeper than the core drill.

While electrical prospecting is thus a competitor of the core drill, it may be used as an aid in making correlations between core drill holes and also between wells in cases where geological correlation is difficult. This is accomplished by making resistivity measurements in the holes at a series of depths and correlating on a resistivity basis.

Electrical methods of geophysical exploration may be divided into two classes, direct current methods and alternating current methods. The first of these also includes the use of very low frequency alternating currents which are commonly used to eliminate the effects of polarization. The frequency, however, is low enough so that the effect of currents induced by the varying magnetic field may be neglected. Induced currents play the major role in the second, or alternating current method of prospecting.

II. Direct Current Methods of Prospecting

There are a great many direct current methods available for the determination of the changes in resistivity with depth, but the same fundamental

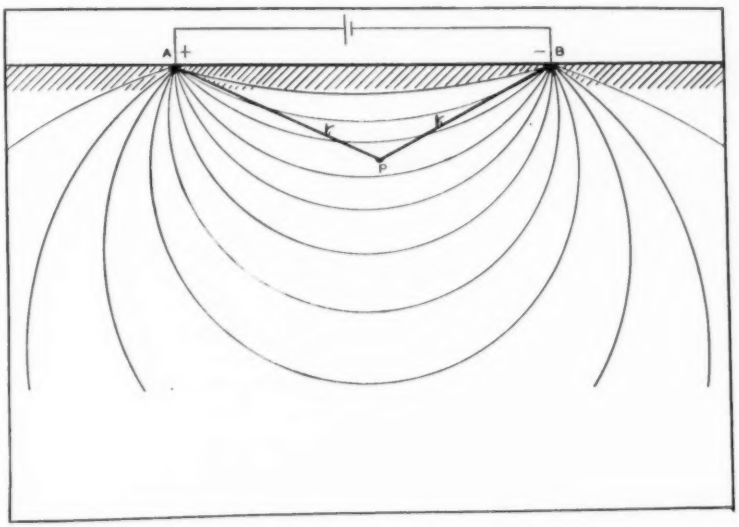

Fig. 1. Lines of flow in a homogeneous earth.

principles apply to all of them. A common practice is to pass the current into the earth by means of two electrodes on the surface, as shown in Fig. 1. If the earth is homogeneous, the lines of flow are arcs of circles connecting the electrodes. The potential on the surface of the earth is then called the normal

potential. Any changes of resistivity in the subsurface beds alter this normal potential. Thus measurements of potential on the surface of the earth give an indication of these changes of resistivity. Most prospecting methods differ only in the spacing of the electrodes and in the manner in which the potential is measured. The same mathematical theory thus applies to all of them.

The theory is simplified by the fact that problems of steady flow in continuous media are the same mathematically as similar electrostatic problems. A current I entering an earth of conductivity σ at an electrode A may be replaced by a charge of $q = I/2\pi\sigma$ at A. The potential at any point in the earth, due to this electrode is $V = I/2\pi\sigma r$, where r is the distance from the electrode.

If several electrodes A_n are on the surface of the earth with currents I_n entering the earth, the potential at any point P in the earth is

$$V = \frac{1}{2\pi\sigma} \sum \frac{I_n}{r_n} \tag{1}$$

where r_n is the distance from P to A_n. It may be noted that algebraic values of the current must be used in this sum. Thus the potential at P in Fig. 1 for the current flow between two electrodes is

$$V = \frac{I}{2\pi\sigma}\left(\frac{1}{r_1} - \frac{1}{r_2}\right). \tag{2}$$

This additive property of the potentials is general and is not confined to a homogeneous earth. Thus we may simplify the theoretical discussion by considering the potentials due to a single electrode.

In order to illustrate the type of departure from normal potential that may be expected, a simple problem will be solved. The earth is of conductivity σ_1, to a depth a and is of conductivity σ_2, below this depth. A current I flows

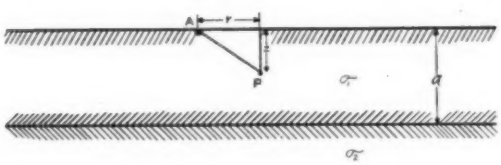

Fig. 2. Coordinate system for two-layer problem.

into the earth at the origin of cylindrical coordinates r, z, ϕ. (See Fig. 2.) Let the potential in the surface layer be V_1 and the potential for $z > a$ be V_2. The conditions to be satisfied by the potential are:

(1) The differential equations expressing the continuity of the currents

$$\nabla^2 V_1 = \frac{\partial^2 V_1}{\partial r^2} + \frac{1}{r}\frac{\partial V_1}{\partial r} + \frac{\partial^2 V_1}{\partial z^2} = 0 \tag{3.1}$$

$$\nabla^2 V_2 = 0. \tag{3.2}$$

(2) The continuity of current and potential at the boundary

$$V_1 = V_2 \text{ at } z = a \tag{3.3}$$

$$\sigma_1 \partial V_1/\partial z = \sigma_2 \partial V_2/\partial z. \tag{3.4}$$

(3) No current flow out of the surface of the earth

$$\partial V_1/\partial z = 0 \text{ at } z = 0. \tag{3.5}$$

(4) The conditions near the electrode

$$V_1 \rightarrow \frac{I}{2\pi\sigma_1(r^2 + z^2)^{1/2}} \text{ as } r_1 z \rightarrow 0. \tag{3.6}$$

(5) The conditions at ∞

$$\begin{aligned} V_1 &\rightarrow 0 \text{ as } \quad r \rightarrow \infty \\ V_2 &\rightarrow 0 \text{ as } r_1 z \rightarrow \infty . \end{aligned} \tag{3.7}$$

This problem has usually been solved by the method of images, which results in an infinite series. An integral solution which is often more useful[1] may be obtained from the following solution of $\nabla^2 V = 0$

$$V = \int_0^\infty [f_1(\lambda)e^{\lambda z} + f_2(\lambda)e^{-\lambda z}]J_0(\lambda r)d\lambda \tag{4}$$

where f_1 and f_2 may be any functions of λ subject to the condition that the integral converges. Thus if

$$V_1 = \frac{I}{2\pi\sigma_1}\int_0^\infty [(A(\lambda) - 1)e^{\lambda z} + A(\lambda)e^{-\lambda z}]J_0(\lambda r)d\lambda \tag{5}$$

and

$$V_2 = \int_0^\infty B(\lambda)e^{-\lambda z}J_0(\lambda r)d\lambda \tag{6}$$

where $A(\lambda)$ and $B(\lambda)$ are arbitrary functions of λ all the conditions (3) will be satisfied except the conditions at the boundary $z = a$. Of course, $A(\lambda)$ and $B(\lambda)$ must be chosen in such a manner that the integrals converge. These functions are determined by substituting V_1 and V_2 given by Eqs. (5) and (6) in the boundary conditions (3.3) and (3.4). The solution for $A(\lambda)$ is

$$A(\lambda) = \frac{1}{1 + ke^{-2\lambda a}} \tag{7}$$

where

$$k = \frac{\sigma_2 - \sigma_1}{\sigma_2 + \sigma_1}.$$

On the surface of the earth, $z = 0$, the potential is

$$V_1 = \frac{I}{2\pi\sigma_1}\int_0^\infty \left[\frac{1 - ke^{-2\lambda a}}{1 + ke^{-2\lambda a}}\right]J_0(\lambda r)d\lambda \tag{8}$$

[1] This method is especially suitable when a number of horizontal layers are present.

$$= \frac{I}{2\pi\sigma_1 r}\int_0^\infty \left[\frac{1 - ke^{-2\beta a/r}}{1 + ke^{-2\beta a/r}}\right] J_0(\beta)d\beta \tag{9}$$

where

$$\beta = \lambda r.$$

When a/r is large, the integral becomes

$$V_1 = \frac{I}{2\pi\sigma_1 r}\int_0^\infty J_0(\beta)d\beta = \frac{I}{2\pi\sigma_1 r} \quad r \ll a \tag{10}$$

which is the normal potential for a homogeneous medium of conductivity σ_1. When a/r is small, the exponentials are approximately equal to unity, and the potential becomes

$$V_1 = \frac{I}{2\pi\sigma_1 r}\,\frac{1-k}{1+k} = \frac{I}{2\pi\sigma_2 r} \quad r \gg a. \tag{11}$$

This is the normal potential for a homogeneous medium of conductivity σ_2.

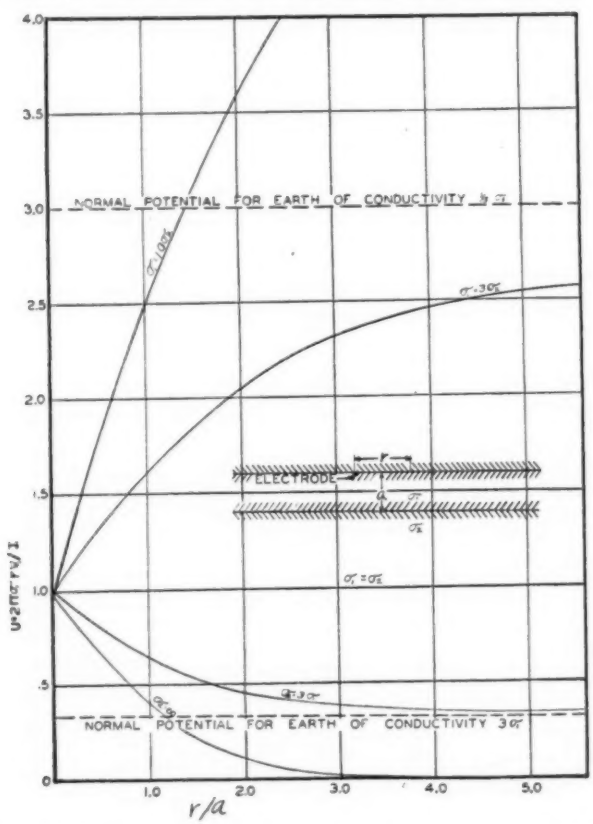

Fig. 3. Theoretical potential curves for two conducting layers.

This example illustrates a fundamental property of this type of prospecting. The potentials at points further and further from the electrode are in-

fluenced more and more by the deeper layers. The surface layers become less important. In order to illustrate this effect values of $u = 2\pi\sigma_1 r V_1 / I$ were plotted against values of r/a for various ratios of σ_2/σ_1 (Fig. 3). The potentials are multiplied by r in order to eliminate the normal drop of potential with distance. If the ground is homogeneous the potential curve is a horizontal line, the ordinate of which depends on the conductivity. Thus for an earth of conductivity σ_1, $u = 1$ for all values of r. For any other conductivity σ_2, $u = \sigma_1/\sigma_2$ for all r. In Fig. 3 the earth is of conductivity σ_1 to a depth a, below which the conductivity is σ_2. The curves for u start at $u = 1$ for $r = 0$ then approach the lines $u = \sigma_1/\sigma_2$ asymptotically. These curves illustrate the gradually increasing effect of the lower medium as r/a increases.

Fig. 4 shows similar curves for a slightly more complicated case. There is a surface layer of conductivity σ_1, then a second layer of conductivity σ_2, and finally a medium whose conductivity is the same as that of the surface layer,

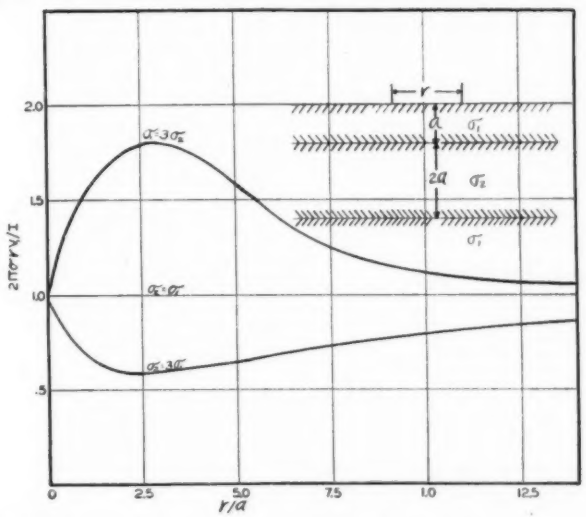

Fig. 4. Theoretical potential curves for three conducting layers.

σ_1. The thickness of the surface layer is a and that of the second layer $2a$. Two curves are shown in Fig. 4: (a) for $\sigma_2 = 3\sigma_1$, and (b) for $\sigma_1 = 3\sigma_2$. In curve (a) $u = 1$ at $r = 0$, and then u decreases for increasing r reaching a minimum at about $r = 2.5a$ where the second layer has most effect and finally approaching the line $u = 1$ asymptotically. Curve (b) rises from $u = 1$ when $r = 0$ to a maximum and then decreases rapidly as it approaches its asymptote $u = 1$. Thus the effect of the second layer on the potential is small near the electrode. It is a maximum at a distance equal to roughly the depth of the layer, and the effect diminishes as r increased beyond this point. These simple cases are sufficient to show how the potential due to a single electrode is affected by vertical changes in conductivity.

In practice it has usually been found more convenient to measure potential differences rather than absolute potential. One method is to put the current electrodes a long distance apart and measure potential differences in

the neighborhood of one electrode so that the potentials due to the distant electrode may be neglected. The potential difference is measured between two search electrodes at distances r and $2r$ from the current electrode. If the earth is homogeneous and of conductivity σ, the potential difference between the search electrodes is:

$$V_r = V(r) - V(2r) = \frac{I}{2\pi\sigma}\left[\frac{1}{r} - \frac{1}{2r}\right] = \frac{I}{4\pi\sigma r}. \tag{12}$$

The resistivity ρ may, therefore, be determined as

$$\rho = 1/\sigma = 4\pi r V_r/I. \tag{13}$$

If the earth is not homogeneous, Eq. (13) may be used to define the apparent resistivity. Let $V_1(r)$ be the potential of the current electrode. Then the apparent resistivity is

$$\rho_a = \frac{4\pi r}{I}\left[V_1(r) - V_1(2r)\right]. \tag{14}$$

This function is not independent of r except in the special case of a homogeneous earth. The variation of the apparent resistivity with r is used to determine the changes of resistivity with depth. As may be expected from the potential variation, the apparent resistance is equal to the resistance of the upper layer when r is small. As r increases the apparent resistance approaches the resistivity of the second layer. Long before it reaches this value, a third layer may affect it, and then the apparent resistivity will have a tendency to approach the resistivity of this layer. When a great many layers are present the resultant curve may be rather complicated, and the interpretation difficult.

A more common method of prospecting which leads to the same value for the apparent resistivity was developed by Wenner and Gish and Rooney. The two current electrodes and the two search electrodes are placed in a line equidistant from one another.

$$\underset{A}{\cdot}\xleftarrow{r}\rightarrow\underset{C}{\cdot}\xleftarrow{r}\rightarrow\underset{D}{\cdot}\xleftarrow{r}\rightarrow\underset{B}{\cdot}$$

The current enters the earth at A and leaves at B. The potential difference is measured between C and D. The common spacing is r. The apparent resistivity is again given by Eq. (14), where V_1 again represents the potential due to a single electrode. The advantages of the method are that the electrode spacing need not be excessively great, and that possible errors due to neglecting part of the field do not enter. However, to obtain values of the apparent resistivity for different values of r, all four electrodes must be moved instead of only two as in the previous method. This method is often called the Gish-Rooney method of electrical prospecting.

In order to illustrate the results that may be expected under favorable conditions, a typical Gish-Rooney survey which was taken along a line of core drill holes is shown in Figs. 5 and 6. Fig. 5 shows the curves of apparent

resistivity at the six stations shown on the profile of Fig. 6. These curves show a layer of very high resistivity which is at the surface at stations *A* and *B* on the west, and then dips to the east. A layer of low resistivity is on the surface at stations *C*, *D*, *E* and *F*, below which there is the high resistivity layer. There is apparently a very thick layer of low resistivity below this high resistivity layer. A rough qualitative interpretation of these curves is shown

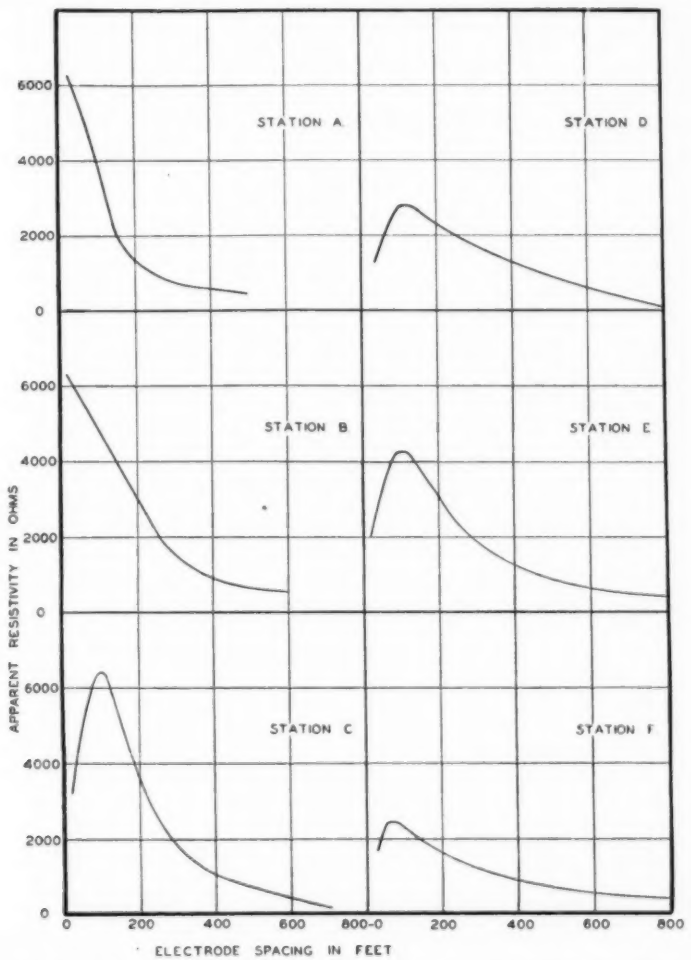

Fig. 5. Apparent resistivity curves obtained by field measurements along profile of Fig. 6.

in Fig. 6. It may be seen that it checks with the core drill holes which show a monocline dipping to the east.

Another method of utilizing the apparent resistivity, the ground resistivity map, has been developed by Schlumberger. An electrode spacing is chosen which is approximately equal to the desired depth of investigation. The apparent resistivities with this electrode spacing are measured at various points on the surface of the earth. A contour map is then made of these values. This method is particularly suitable in a country with a shallow overburden which

covers a thick sediment with a large change in resistivity. Then the contours will roughly approximate the topographical contours on the sediment. Suppose that the sediment has a low resistivity. Then when it is near the surface, the apparent resistivity will be low. When the sediment is deeper, the apparent resistivity will be higher.

The apparent resistivity map, of course, gives only a qualitative interpretation of the electrical data. The methods at present in use for interpreting Gish-Rooney surveys give only approximate depths to the points at which sharp changes occur in the resistivity of the earth. These approxima-

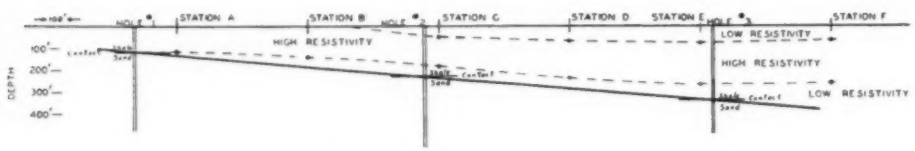

Fig. 6. Qualitative interpretation of Gish-Rooney survey shown in Fig. 5, giving correlation with core drill.

tions become worse and worse as the number of resistivity contrasts become greater. A rigorous solution of the problem of finding the variation of resistivity with depth from surface potential measurements under the assumption that the resistivity is substantially a function of the depth only could be based upon solutions of the potential problem given by the theory of images. However, these solutions are cumbersome to handle when the number of layers which must be dealt with exceeds two. The integral solution given in this paper is much easier to handle in the case of multilayer problems and should aid in placing the interpretation of direct current surveys on a more exact basis.

III. Electromagnetic Methods

Electromagnetic methods depend primarily on the induction of currents in the earth by means of a primary current on the surface. Inhomogeneities in the earth will affect the resultant field on the surface. In oil prospecting these inhomogeneities take the form of practically horizontal beds with different conductivities. Many different schemes have been used with varying success. The factors which should be taken into consideration are:

(1) The primary field should be as simple as possible so that the associated theoretical problems may be solved.

(2) The frequency should be chosen so that the inhomogeneities which are being followed will have a maximum measurable effect on the field.

(3) The field components should be measured at points where the beds produce the greatest measurable anomaly in the field components. Thus both the magnitude of the change and the percentage change in the field should be considered.

The quantities which may be measured are the magnitude and time phase of the components of the electric and magnetic vectors on the surface of the earth.

There are two simple primary fields for which the theoretical problems have been solved; the field of a circular coil, which may be considered to be an oscillating magnetic dipole with its axis perpendicular to the plane of the coil and the field of a long horizontal wire. Most electromagnetic methods have made use of one of these.

The basic equations that govern alternating currents are Maxwell's equations. The field components must satisfy the wave equation:[2]

$$\nabla^2 A_1 - \mu\sigma \frac{\partial A_1}{\partial t} - \mu p \frac{\partial^2 A_1}{\partial t^2} = 0. \tag{15}$$

If the impressed forces are simple harmonic functions of time, A_1 can be represented as

$$A_1 = \text{real part of } Ae^{-i\omega t}$$

where A is in general complex. Then Eq. (15) becomes

$$\nabla^2 A + i\omega\mu A(\sigma - i\omega p) = 0. \tag{16}$$

This equation shows the relative importance of the conductivity and the relative permittivity or dielectric constant in determining the current flow in the earth. If $\sigma \gg \omega p$ the conductivity will be the determining factor. If $\omega p \gg \sigma$ the dielectric constant will be the determining factor. If the conductivity is 10^{-4}ohms·cm and the dielectric constant is 10, then the two are of equal importance when the frequency is

$$f = \frac{10^{-4}}{2\pi \times 10 \times 8.85 \times 10^{-14}} = 1.8 \times 10^7.$$

It will be shown later, however, that the high frequency waves are very highly damped by the surface layers, and are thus not important in oil prospecting. At the lower frequencies, the conductivity governs the distribution of current in the earth. Methods of locating oil directly by means of its dielectric constant are thus not very hopeful.

In order to show the importance of using a low frequency, a particularly simple situation has been assumed. The primary field is due to a vertical oscillating electric dipole on the surface of the earth.[3] The earth is homogeneous and of conductivity σ to a depth h, below which there is a layer of very high conductivity which we shall assume to be infinite. The problem to be solved is—at what frequency will the highly conducting layer have the biggest influence on the field at the surface of the earth?

The field may be conveniently represented by the Hertzian function Π, which satisfies the wave equation

[2] The rationalized practical system of units is used. Magnetic intensity H is expressed in ampere turns per cm. Magnetic flux density B is expressed in webers per cm^2 ($=10^8$ Maxwell's per cm^2). Permeability $\mu = B/H = 4\pi \times 10^{-9}$ for free space. Electric intensity E is expressed in volts per cm. Displacement D is expressed in coulombs per cm^2. Permittivity $\rho = D/E = 8.85 \cdot 10^{-14}$ for free space. Conductivity σ is expressed in mhos per cm.

[3] Riemann-Webers, Diff. Gleich. der Phys. (Seventh ed. 1925) p. 542.

$$\nabla^2\Pi + k_1^2\Pi = 0 \begin{cases} k_1^2 = k_0^2 = \mu p_0 \omega^2 = \omega^2/c^2 \text{ in the air} \\ k_1^2 = k^2 = \mu p \omega^2 + i\mu\sigma\omega \text{ in the earth} \end{cases} \tag{17}$$

where c is the velocity of light.

Cylindrical coordinates (r, z, ϕ) with the origin at the dipole will be used. The electric intensity E and the magnetic intensity H are:

$$\begin{aligned} E_\phi &= H_z = H_r = 0 \\ E_r &= \partial^2\Pi/\partial r \partial z \\ E_z &= -\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \Pi}{\partial r}\right) \\ H_\phi &= \frac{ik_1^2}{\mu\omega}\frac{\partial \Pi}{\partial r}. \end{aligned} \tag{18}$$

We shall take the Hertzian function for the air to be Π_0, and the Hertzian function for the earth to be Π. The boundary conditions then are

$$\left.\begin{aligned} k_0^2\Pi_0 &= k^2\pi \\ \partial\Pi_0/\partial z &= \partial\pi/\partial z \end{aligned}\right\} \text{ at } z = 0 \tag{19}$$

and

$$\frac{\partial \Pi}{\partial z} = 0 \text{ at } z = -h. \tag{20}$$

The Hertzian function for a dipole in free space is

$$\Pi_d = \int_0^\infty \frac{\lambda}{(\lambda^2 - k_0^2)^{1/2}} e^{-z(\lambda^2 - k_0^2)^{1/2}} J_0(\lambda r) d\lambda. \tag{21}$$

We, therefore, take for Π_0

$$\Pi_0 = \int_0^\infty \left[\frac{\lambda}{(\lambda^2 - k_0^2)^{1/2}} - f(\lambda)\right] e^{-z(\lambda^2 - k_0^2)^{1/2}} J_0(\lambda r) d\lambda \quad z > 0 \tag{22}$$

and for Π

$$\Pi = \int_0^\infty [g_1(\lambda) e^{z(\lambda^2 - k^2)^{1/2}} + g_2(\lambda) e^{-z(\lambda^2 - k^2)^{1/2}}] J_0(\lambda r) d\lambda. \tag{23}$$

The functions f, g_1, and g_2 are chosen so as to satisfy the boundary conditions Eqs. (19) and (20). The solution for Π_0 is

$$\Pi_0 = \int_0^\infty \frac{2k^2\lambda e^{-z(\lambda^2 - k_0^2)^{1/2}} J_0(\lambda r) d\lambda}{k^2(\lambda^2 - k_0^2)^{1/2} + k_0^2(\lambda^2 - k^2)^{1/2} \tanh h(\lambda^2 - k^2)^{1/2}}. \tag{24}$$

The normal field for a homogeneous earth of conductivity σ is

$$\Pi_n = \int_0^\infty \frac{2k^2\lambda e^{-z(\lambda^2 - k_0^2)^{1/2}} J_0(\lambda r) d\lambda}{k^2(\lambda^2 - k_0^2)^{1/2} + k_0^2(\lambda^2 - k^2)^{1/2}}. \tag{25}$$

The change in the field due to the layer of high conductivity is:

$$\Pi_0 - \Pi_n = \int_0^\infty \phi(\lambda) e^{-z(\lambda^2-k_0^2)^{1/2}} \lambda J_0(\lambda r) d\lambda \tag{26}$$

where

$$\phi(\lambda) = \frac{2k^2k_0^2(\lambda^2-k^2)(1-\tanh h(\lambda^2-k^2)^{1/2})}{[k^2(\lambda^2-k_0^2)^{1/2}+k_0^2(\lambda^2-k^2)^{1/2}][k^2(\lambda^2-k_0^2)^{1/2}+k_0^2(\lambda^2-k^2)^{1/2}\tanh h(\lambda^2-k^2)^{1/2}]} . \tag{27}$$

In order to calculate the effect of the layer on the normal field, one could calculate the components of this change in field and compare them with the normal field components. A simpler way to compare them is to compare the power flow associated with the change in field to the normal power flow. The changes in the field components are:

$$\begin{aligned} (E_r)_0 - (E_r)_n &= \int_0^\infty \phi(\lambda)(\lambda^2 - k_0^2)^{1/2} e^{-z(\lambda^2-k_0^2)^{1/2}} \lambda^2 J_1(\lambda r) d\lambda \\ (H_\phi)_0 - (H_\phi)_n &= \frac{ik_0^2}{\mu\omega} \int_0^\infty \phi(\lambda) e^{-z(\lambda^2-k_0^2)^{1/2}} \lambda^2 J_1(\lambda r) d\lambda . \end{aligned} \tag{28}$$

The power flow of this field through a horizontal plane in the air is:

$$\begin{aligned} P &= \pi \int_0^\infty [(E_r)_0 - (E_r)_n][(H_\phi)_0 - (H_\phi)_n]^* r dr \\ &= \frac{\pi k_0^2}{\mu\omega} \int_0^{k_0} (k_0^2 - \lambda^2)^{1/2} \left| \phi(\lambda) \right|^2 \lambda^3 d\lambda . \end{aligned} \tag{30}$$

The asterisk indicates that the complex conjugate of $(H_\phi)_0 - (H_\phi)_n$ is to be taken.

For all but very high frequencies, k^2 is very much greater than k_0^2. Under this condition, the power flow is approximately

$$P = \frac{4\pi k_0^6[(1-R)^2 + I^2]}{(2)^{1/2}\mu\omega \left| k \right| (1 - R + I)} \log \frac{1}{R - I} \tag{31}$$

where R and I are real, and

$$\tanh h \left| k \right| \frac{1+i}{2^{1/2}} = R + iI . \tag{32}$$

The normal power flow from a homogeneous earth of conductivity σ is

$$P_n = \frac{4\pi k_0^2 \left| k \right|^4}{\mu\omega} \int_0^{k_0} \frac{(k_0^2 - \lambda^2)^{1/2}\lambda^3 d\lambda}{\left| k^2(\lambda^2 - k_0^2)^{1/2} + k_0^2(\lambda^2 - k^2)^{1/2} \right|^2} \tag{33}$$

which becomes for $k^2 \gg k_0^2$

$$P_n = 8\pi k_0^5/3\mu\omega . \tag{34}$$

The ratio of the power flow of the change in field to the normal power flow is:

$$P/P_n = \frac{3}{2} \frac{k_0}{2^{1/2} \mid k \mid u} g(u) \frac{2.81 \times 10^{-3}}{h\sigma} g(u) \tag{35}$$

where

$$g(u) = u \frac{(1-R)^2 + I^2}{1 - R + I} \log \frac{1}{R - I} \tag{36}$$

$$u = h(\mu\sigma\omega)^{1/2} \tag{37}$$

$$\tanh u \frac{1+i}{2^{1/2}} = R + iI. \tag{38}$$

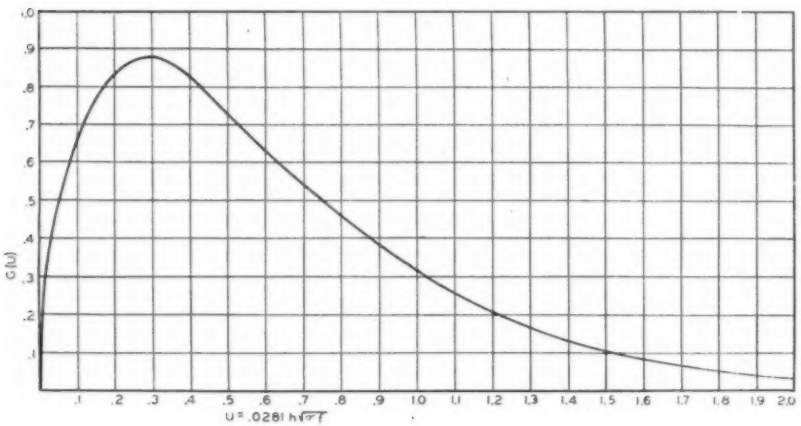

Fig. 7. Power ratio curve for vertical electric dipole.

Fig. 7 gives a graph of $g(u)$ which shows a maximum at $u = 0.29$. This gives the optimum frequency in terms of the depth of investigation and the conductivity of the upper beds. This frequency is

$$f = \frac{10^9 u^2}{8\pi^2 h^2 \sigma} = \frac{(0.29)^2 \cdot 10^9}{8\pi^2 h^2 \sigma} = \frac{10^7}{h_m{}^2 \sigma} \tag{39}$$

if h is measured in meters. If $\sigma = 10^{-4}$ and $h = 100$ meters, the best frequency to use is about 100 cycles. If $h = 300$ meters the optimum is about 10 cycles. At higher frequencies the effect of the conducting layer on the field becomes smaller. Thus for $\sigma = 10^{-4}$ and $h = 100$ meters, the effect of the layer with a frequency of 2500 cycles is only about one tenth of the effect at 100 cycles. At very low frequencies only a small percentage of the power input enters the ground. At high frequencies, most of the power that enters the ground is absorbed by the surface layers.

A similar investigation has been carried out to determine the effect of a highly conducting layer on the field of a long horizontal wire carrying an alternating current. As in the previous case, the earth is assumed homogeneous with conductivity σ to a depth h, below which the conductivity is infinite.

The wire is on the z-axis of rectangular coordinates, the y-axis is vertical and the x-axis is horizontal and perpendicular to the wire. Maxwell's equations for the conducting medium reduce to:

$$\begin{aligned} E_x = E_y = H_z &= 0 \\ \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} &= (\sigma - i\omega p)E \\ \frac{\partial E_z}{\partial y} &= i\omega\mu H_x \\ \frac{\partial E_z}{\partial x} &= -i\omega\mu H_y \end{aligned} \tag{40}$$

or

$$\nabla^2 E_z + k^2 E_z = 0.$$

The equations for the air may be obtained by setting $\sigma = 0$ and $p = p_0$. Let us take

$$\begin{aligned} E_z &= E_0 \text{ in the air} \\ E_z &= E \text{ in the earth.} \end{aligned} \tag{41}$$

The boundary conditions then are

$$\left.\begin{aligned} E_0 &= E \\ \partial E_0/\partial y &= \partial E/\partial y \end{aligned}\right\} \text{ at } y = 0 \tag{42}$$

$$E = 0 \text{ at } y = -h \tag{43}$$

and

$$E_0 \to k_0 \log (x^2 + y^2)^{1/2} \text{ as } x, y \to 0. \tag{44}$$

The solution of these equations for E_0 is

$$E_0 = k_0 \int_0^\infty \frac{2e^{-y(\lambda^2 - k_0^2)^{1/2}} \cos \lambda x d\lambda}{(\lambda^2 - k_0^2)^{1/2} + (\lambda^2 - k^2)^{1/2} \coth h(\lambda^2 - k^2)^{1/2}}. \tag{45}$$

The normal field for a homogeneous earth is

$$E_n = k_0 \int_0^\infty \frac{2e^{-y(\lambda^2 - k^2)^{1/2}} \cos \lambda x d\lambda}{(\lambda^2 - k_0^2)^{1/2} + (\lambda^2 - k^2)^{1/2}}. \tag{46}$$

Thus the change in the field due to the conducting layer is

$$E_0 - E_n = k_0 \int_0^\infty \phi(\lambda) e^{-y(\lambda^2 - k_0^2)^{1/2}} \cos \lambda x d\lambda \tag{47}$$

where

$$\phi(\lambda) = \frac{2(\lambda^2 - k^2)^{1/2}(1 - \coth h(\lambda^2 - k^2)^{1/2})}{[(\lambda^2 - k_0^2)^{1/2} + (\lambda^2 - k^2)^{1/2}][(\lambda^2 - k_0^2)^{1/2} + (\lambda^2 - k^2)^{1/2}\coth h(\lambda^2 - k^2)^{1/2}]} \quad (48)$$

and the change in magnetic intensity is:

$$(H_x)_0 - (H_x)_n = \frac{ik_0}{\mu\omega}\int_0^\infty (\lambda^2 - k_0)^2\phi(\lambda)e^{-y(\lambda^2 - k_0^2)^{1/2}}\cos \lambda x d\lambda. \quad (49)$$

The power flow through a horizontal plane in the air per centimeter of wire length is

$$P = \tfrac{1}{2}\int_\infty^\infty [E_0 - E_n][(H_x)_0 - (H_x)_n]^* dx = \frac{k_0^2\pi}{4\mu\omega}\int_0^{k_0} |\phi(\lambda)|^2(k_0^2 - \lambda^2)^{1/2}d\lambda. \quad (50)$$

When $k^2 \gg k_0^2$ this integral reduces to

$$P = \frac{\pi^2 k_0^4}{4\mu\omega |k|^2}[(1 - R)^2 + I^2] \quad (51)$$

where R and I are real and

$$R + iI = \tanh h|k|\left(\frac{1 + i}{2^{1/2}}\right). \quad (52)$$

The normal power flow through the plane for a homogeneous earth is:

$$P = \frac{\pi^2 k_0^4}{4\mu\omega |k^2|}. \quad (53)$$

As the frequency approaches zero, the ratio $k_0^4/\omega|k|^2$ approaches zero, so that both the normal power flow and the power flow of the change in field due to the layer approach zero. The ratio, however, is

$$P/P_n = (1 - R)^2 + I^2 \quad (54)$$

where

$$R + iI = \tanh u \frac{(1 + i)}{(2)^{1/2}}$$

$$u = h|k| = 0.0281\, h_{\text{meter}}(\sigma f)^{1/2}.$$

The ratio p/p_n is shown graphically in Fig. 8. It is a maximum at zero frequency, and then drops off as the frequency increases, approaching zero at very high frequencies. In this case the optimum frequency depends primarily on the measuring apparatus. The frequency should be chosen as low as the apparatus permits the field quantities to be measured accurately.

The curve in Fig. (8) shows the absorption of the surface layer. If the conductivity of the surface layer is 10^{-4} ohms and the depth to the conducting layer is 100 meters, the effect of the layer at a frequency of 3000 cycles is only

about one twentieth of the effect at 100 cycles. This curve illustrates the importance of using a low frequency alternating current.

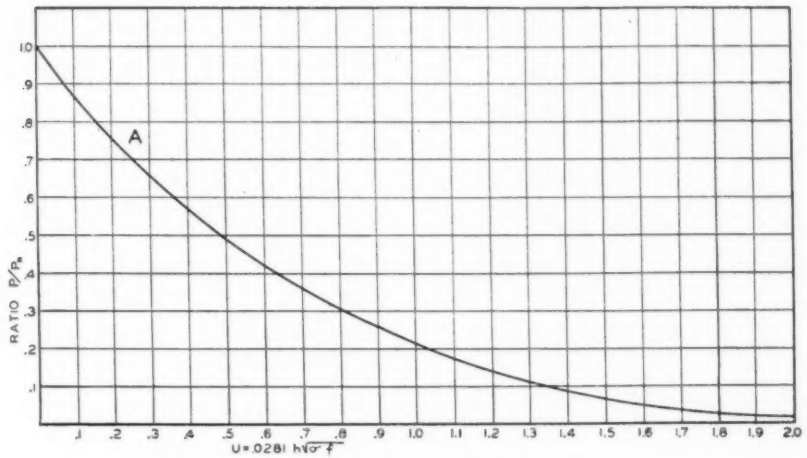

Fig. 8. Power ratio curves for long line.

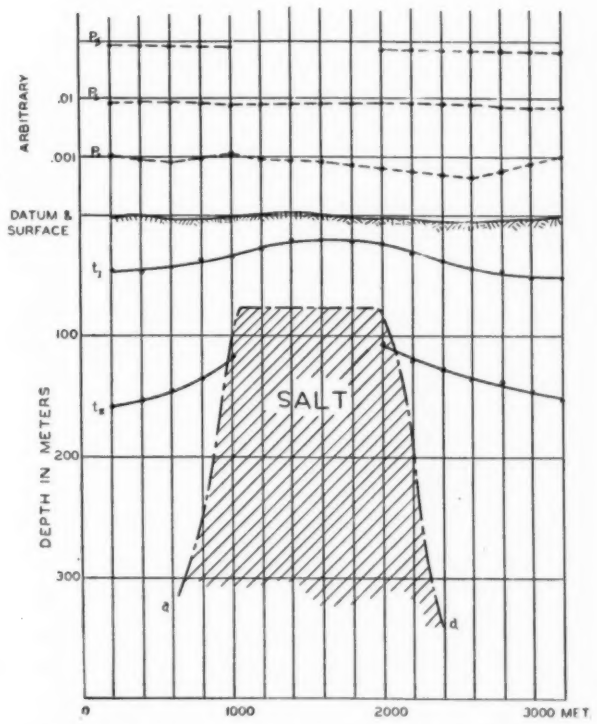

Fig. 9. Electromagnetic survey across a salt dome.

In order to illustrate the results that may be obtained in the field by a good electrical survey under favorable conditions, two surveys made by the Swedish American Prospecting Corporation for the Gulf Oil Corporation will

be described. The method of survey used was the Sundberg method,[4] which will be briefly described.

The primary field is the field of a long horizontal wire carrying an alternating current with a frequency of 100 to 500 cycles. Measurements of the direction and time phase of the magnetic intensity vector are made at a series of points along transverse lines. These measurements are made by means of a search coil connected to a compensating arrangement.

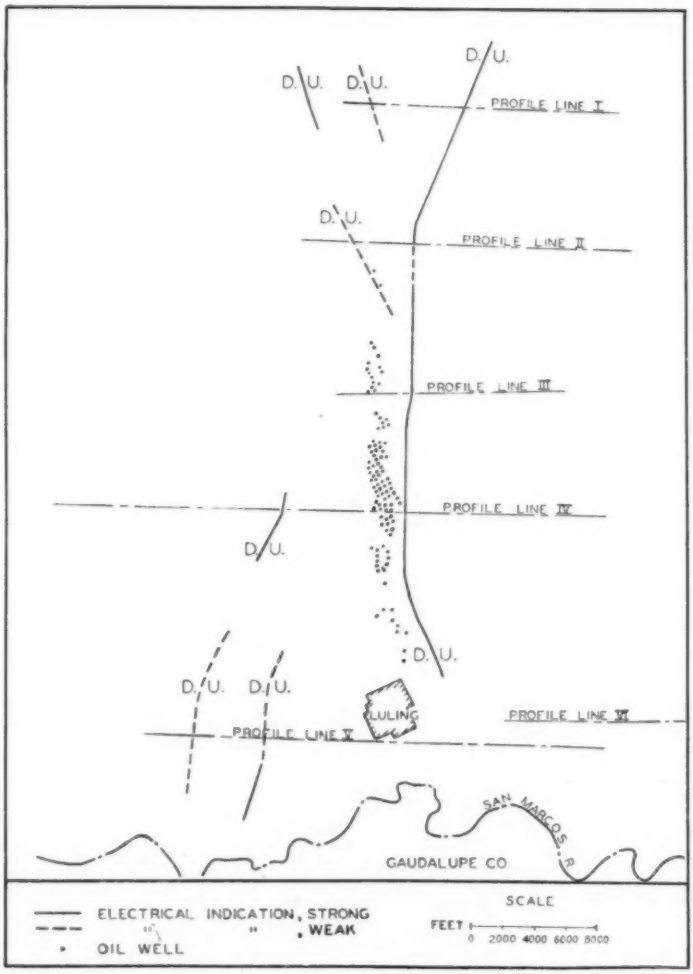

Fig. 10. Fault traced by electrical methods where only one well was drilled in the field.

The induced currents are assumed to be confined to a series of beds which have a much higher conductivity than the surrounding beds. Each of these beds is characterized by a certain induction factor p which depends on its conductivity and thickness, and also on the frequency of the exciting current. From the measurements made at the surface, the induction factors and the

[4] T. Zuschlag, Tech. Pub. No. 313 Inst. of Min. and Met. Eng.

depth of the beds are determined. The conducting beds which are followed are called marker beds.

Fig. 9[5] shows the results of a survey across a shallow salt dome. Three conducting beds were assumed, one at the surface, one at an intermediate depth, and the last at about 400 feet. The intermediate bed is arched over the dome. The deeper conductor is missing where it is pinched out by the salt, and locates the flanks of the dome.

Fig. 10 gives the electrical indication of the Bruner fault zone, Texas. A clay bed at about 500 feet was used as a marker. When the survey was started there was only one producing well in the area. Later production has followed parallel to the electrically indicated fault. The electrical picture was not only qualitively right, but later drilling showed a constant interval between the clay bed and the Austin chalk, which is the producing horizon. Thus, a very good shallow picture of this fault was obtained electrically.

IV. Conclusion

Up to the present point we have discussed in a general way the problems encountered in electrical prospecting under very simple geological conditions. Experience in the field has shown that in general good marker beds can be found within the depth to which investigation can be carried electrically. However, in countries where the shallow sediments have been laid down in the form of lenses, these beds may not be continuous enough to give a structure picture.

On monoclinal slopes where successive beds are outcropping so that the surface material is changing in character, the lateral variation in conductivity of the surface material is large. This resistance variation of the near surface material makes the problem of interpreting electrical measurements somewhat difficult. A condition of this kind was encountered in certain parts of the Balcones fault zone of Texas. The variation in surface resistivity resulted in many apparent fault indications and it was difficult to distinguish these apparent indications from the true ones.

The depth to which geophysical investigations can be carried is of considerable importance in regions where the lower beds which contain the oil are not laid down conformably with the shallow beds. If the beds in which oil must be sought lie below an uncomformity which occurs at a depth of 5000 feet, then it is obvious that a prospecting scheme cannot be used which does not give information below 5000 feet.

Many extravagant claims have been made for the possible depths to which it is possible to work by electrical methods. In direct current cases, if the earth were homogeneous except for two or three sharp breaks in conductivity, it would be possible to map to a large depth. However, even in this ideal case as the depth of investigation is increased, there would be a gradual loss in detail. Actually conditions are much more complicated. As the depth of investigation is increased, more and more discontinuities enter both later-

[5] Surveys shown by Figs. 9 and 10 have previously been published by Zuschlag in A.I.M.M.E. Technical Publication No. 313.

ally and vertically. Consequently, the difficulties of interpreting the measurements become more and more difficult. Up to the present time information of immediate value in oil geology has not been obtained below a depth of about 2000 feet.

In the alternating current case there is a definite depth of current penetration for a given frequency and a given earth resistivity. In attempting to work at deeper and deeper depths, the operating frequency must be lowered continually. In addition to the difficulties mentioned in the direct current case, one is confronted with the problem of measuring very weak electric and magnetic fields of low frequency. In some cases where the resistivity of the surface materials is high the alternating current field is less influenced by lateral irregularities in the surface layer than the direct current field. In the present state of the art, it would take exceptionally favorable conditions to obtain reliable information at a depth much in excess of 1500 feet.

At the present time electrical methods of prospecting for oil seem to be in disrepute. This is partly due to cost of electrical surveys as compared with other geophysical methods and partly due to the failure of the extravagant claims made for the process to materialize. However, the electrical method of prospecting for oil cannot be forgotten because it is one of two prominent geophysical methods in which it is possible to control the field being employed. Improvements in methods of interpretation and in field technique should give electrical methods a definite field of usefulness in prospecting for oil.

A NEW INSTRUMENT FOR MEASURING VERY SMALL DIFFERENCES IN GRAVITY

By Kenneth Hartley
Hartley Gravity Balance Corporation, Houston, Texas

(Received January 2, 1932)

Abstract

Description of a new portable instrument for measuring relative values of gravity to within two or three parts in ten million, specially designed for geophysical exploration. The principles of the design are analysed and the methods for eliminating effects of elastic hysteresis, temperature changes, variations in barometric pressure, etc., are discussed. Also effects of initial stresses in materials, defects in alignment of locking mechanism, inaccurate leveling, etc. These difficulties are serious but seem to have been overcome. Results of preliminary field measurements near Houston, Texas, are presented and some comparison between the type of information given by this instrument and by the torsion balance.

UNTIL now the only successful method of measuring gravity has been with some form of pendulum, in other words determining the acceleration produced, or, we may say, obtaining the value of gravity in terms of inertia. There is only one other possible method and that is to measure it in terms of the elasticity of a spring which can be calibrated in some way. This has the advantage of eliminating time measurements but introduces some other difficulties that are avoided by the pendulum.

In all previous attempts that I know of to use a spring the effort has been to obtain extreme sensitivity by some arrangement approaching unstable equilibrium, and one reason for their failure is that this unstable condition also exaggerates all the causes of error and uncertainty in the indications of the apparatus. In the new instrument the effort has been in the other direction, that is to obtain the maximum stability, the small displacements are then amplified by optical means with less uncertainty.

The fundamental idea of the design is that if we can apply to a suspended mass an upward force almost, but not quite, equal to the pull of gravity and maintain this force *constant* then it will be easy to measure changes in the small additional force required to hold the system in equilibrium. The first problem then is to apply a force that can be maintained constant to the required degree of accuracy. The value of such an instrument for geophysical prospecting depends on its ability to measure very minute differences in the gravitational force, so the instrument has been designed to have the greatest sensitivity that can be of any practical use. Since the change in the gravitational attraction due to a difference of one foot in elevation is almost exactly one part in ten million this was taken as the maximum sensitivity that could possibly be useful; for most purposes three or four parts in ten million will be

all that is desired. This is about ten times the accuracy of the best work done with the pendulum apparatus of the United States Geodetic Survey.

Figure 1 is a simplified diagram of the essential parts of the apparatus. The main spring, which carries about 99.9 percent of the load, is formed from heavy wire and not stressed to more than 20 percent of its apparent elastic limit, so as to give the greatest possible dependability. A very light beam, hinged at the left side, is flexibly attached between the spring and the weight

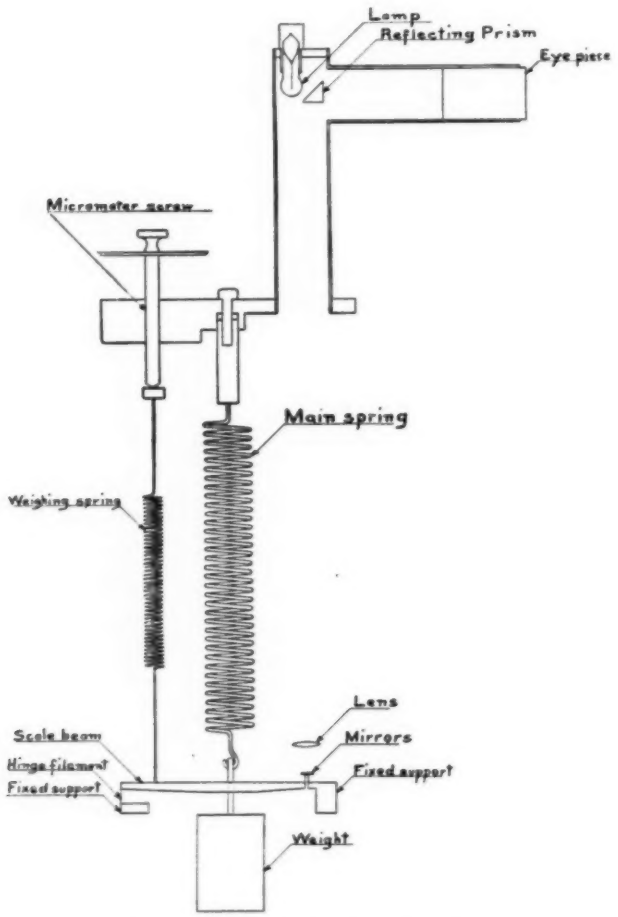

Fig. 1. Schematic diagram of gravimeter.

and carries two small mirrors at the free end. These mirrors act on the principle of the optical lever and are so mounted that a vertical displacement of the beam causes them to turn in opposite directions. This doubles the amplification but a much more important fact is that it avoids the necessity for maintaining a fixed point of reference in the telescope; when the two mirrors are in the same plane the two images of the lamp filament as seen in the eye-piece will be in one straight line no matter how much the eye-piece and other optical parts may be displaced. A very light weighing spring is attached to the

beam between the center and the hinged end and the tension in this spring is controlled by a micrometer screw above.

If the instrument is adjusted at a certain location so that the two images of the lamp filament are in line and it is then taken to a different location, the difference in gravity will cause a slight displacement of the images which are then brought back to the zero position by altering the tension in the weighing spring; the difference between the two readings of the micrometer dial is then a measure of the difference in the force of gravity at the two places.

The difficulties involved in carrying out this idea to the necessary degree of precision have been greatly increased by the requirement that the instrument be rugged enough to stand transportation over rough roads and small and light enough to be carried by two men where there are no roads.

Suppose a mass of 100 grams or more suspended by a spring of such stiffness that the total elongation, from no load to full load, is 100 millimeters; then, to maintain the tension constant to one part in ten million would require; first, that the length be controlled to within one one-hundred-thousandth of a millimeter; second, that the temperature be maintained constant for the whole period required for a series of observations so that the elongation of the spring due to temperature effect minus the linear expansion of the supporting frame will be within the same limit; third, that effects of elastic hysteresis be eliminated not only from the main spring but from the flexible supports of the beam and mirrors also.

In the present instrument the distance between the thin ribbons supporting the mirrors is one millimeter and the distance from the mirrors to the eye-piece is 600 millimeters, so that the amplification for each mirror is 1200 times, but as the two mirrors move in opposite directions the relative displacement of the images is amplified 2400 times. As the beam ratio is 1.8 the total amplification is 4320, but as this is observed through an eye-piece with a magnification of 14 the apparent amplification is over 60,000, so that a movement of the suspended mass of one one-hundred-thousandth of a millimeter will look like six-tenths of a millimeter and be easily observed. If the force of gravity is increased by one part in ten million, part of the increase is used in stretching the main spring and part in bending the small filaments that support the beam and mirrors, so the sensitivity in terms of force is not so great as this, but the first requirement is accomplished—the maintainance of the length of the main spring to within one part in ten million.

I expected to be able to make the beam and mirror system so flexible that the resistance to movement would not be any greater than that of the main spring, but I have not succeeded in doing so. Although the filaments that carry the mirrors are only 0.01 mm thick, it is the stiffness of these filaments that limits the sensitivity of the instrument. The present instrument is satisfactory for readings of one part per million but is difficult to read any closer than that so a new beam system is being made which is expected to be at least four times as sensitive as the present one.

The temperature effect on the spring is very large and unless accurately controlled will completely mask the small variations in force that are to be

measured. The ideal material for the spring would be a nonmagnetic metal with a high elastic limit but a low value for Young's modulus and a low temperature coefficient. This combination might be found in one of the less common metals or alloys but information on the elastic properties of these metals, particularly temperature effects, was not to be had. Assuming that the temperature coefficient of elasticity would probably be smaller in a metal with very high melting point led to the consideration of tantalum, molybdenum, etc., and several springs were made of approximately the dimensions that would be required and tested in that form. The final choice was an alloy of tantalum and tungsten which seems to be fairly satisfactory. The spring is wound with a high initial tension so that when carrying its full load the coils are separated only about half a millimeter, making the total length of the stretched spring less than ten centimeters. The frame of the instrument is made of an aluminum alloy, which has the highest coefficient of expansion of any suitable metal, and is constructed on the principle of the old grid-iron compensated pendulum so that the total length of the high expansion members is about 80 centimeters, the low expansion rods are invar steel. This compensation is mainly for the purpose of avoiding too great a change in the relative position of the parts when the working temperature is altered. In order that the whole instrument be in a condition of thermal equilibrium and free from temperature stresses it must be kept at a constant temperature day and night during a whole series of observations; it is therefore enclosed in a new type of portable thermostat, designed on the principle worked out at the Bureau of Standards two or three years ago, which controls the temperature to within one ten-thousandth of a degree. The heating is electrical and is designed to work with a six volt storage battery.

A study of the action of this instrument while the temperature is either rising or falling will quickly convince any one that any plan for making a temperature correction, instead of maintaining a constant temperature, is wholly out of the question. In the first place, even with the best possible compensation the temperature would have to be measured to within one hundredth of a degree, and in the second place, it is impossible to have the entire apparatus at the same temperature unless it is kept constant for a considerable time. In the present instrument if the temperature is altered by two or three degrees it requires at least five hours to reach a condition of equilibrium and if the change is much greater it may take all day, but with thermostat control the temperature effects are also brought within one part in ten million.

The change in density of the air, due to varying barometric pressure, may alter the buoyant force on the suspended mass by several times the quantity that we are attempting to measure, so the case is closed air-tight and the two operating rods are surrounded by a new type of mercury seal which prevents the passage of any air in or out even with a difference of pressure of more than two inches of mercury.

The effects of elastic hysteresis are eliminated by locking the weight and spring in the zero position between observations, so that the tension on the spring is never varied by more than the difference between the force of gravity

at two successive stations. Any hysteresis effect is then only a small fraction of this amount. This sounds simple but it is difficult to carry out for it requires that the suspended mass be locked to within less than one thousandth of a millimeter of the true position. To meet this requirement and to avoid any strain on any part of the spring from jolts in transportation has necessitated a very elaborated locking mechanism.

In connection with this is another matter that no one seems to think of unless he has been working with a seismograph; that is vibration. The ground is constantly vibrating, and a weight hung on a spring is an excellent seismograph if its motions can be closely observed, so some way must be found to reduce the sensitivity to vibrations while maintaining the sensitivity to changes in the static force. In the present instrument this is accomplished to a certain extent by excessive air damping but this is only satisfactory under very favorable conditions so some modifications are being planned for the beam system which will give a more favorable ratio between these two effects. In the city traffic vibration is the chief difficulty and in the country it is the wind in the trees, so it is impossible to use the instrument in very windy weather but with a light wind readings can be made by watching for a favorable moment, the mirrors may be almost stationary for several seconds at a time and an experienced operator can utilize these moments to secure readings to about one part in a million.

One more difficulty encountered in the construction of the instrument was due to initial strains in the frame. Any change in temperature would alter the elastic forces present and cause the frame to twist enough to throw the beam system badly out of adjustment. To avoid this as far as possible the beam and mirror system is carried by a separate frame entirely independent of that which carries the rest of the mechanism. This frame is made of annealed aluminum tubes very accurately fitted into sockets in the top and bottom plates, and clamped in position by a method designed to avoid putting any strain in the tubes when they are tightened up.

It was found, in the first experimental instrument, that the temperature stresses together with stresses due to imperfect construction of the locking mechanism would pull the beam and mirrors slightly out of alignment when it was locked. This caused distortion of the filament supports of the mirrors and of a light guide spring that was used in that instrument and when released the filaments would not immediately return to their relaxed condition but there would be a slow drift of the readings extending over a period of from half an hour up to several hours before the correct reading was obtained. In the latest instrument this effect is almost eliminated but it is sometimes necessary to wait about ten minutes.

These effects due to elastic hysteresis have been the most serious difficulties encountered, and there is still room for improvement in this direction, but measurements are now being made than can be checked, day after day, to $\pm 1 \cdot 10^{-6}$ and, as soon as a small alteration in the mirror system can be completed, which will give a better ratio between the effect of gravity and the effect of vibrations, I expect to get an accuracy four or five times as great.

The instrument must be very accurately leveled, that is, the axis of the instrument must be in the direction of the gravitational force. An angle of ninety seconds between the axis of the instrument and the true vertical would

Fig. 2. The gravimeter controls and adjustments.

theoretically cause an error of one part in ten million, but it would cause a much greater difference in the actual measurement because of lateral strains in the filaments carrying the beam and mirrors. The present instrument is

Fig. 3. The gravimeter in the field.

easily adjusted to thirty seconds and that seems to be sufficiently close for the purpose.

The dial readings give a direct measure of the differences in gravity and do not require any correction except to multiply by the calibration factor for

the instrument and subtract the effect of elevation, which is one part in ten million for every foot. These differences in elevation can be determined to within three or four feet with the Paulin Altimeter, provided that it is used according to a particular system that has been worked out for this purpose, but if greater accuracy is desired the land must be surveyed. These facts make the use of the gravity balance much simpler than any other geophysical instrument except the magnetometer.

Recent field work has shown that in reasonably good weather it is easy to make three stations per hour. The torsion balance makes three stations per day and the Government pendulum apparatus three or four stations per month. This measurement which takes only a few minutes appears to have four or five times the accuracy of a single run of the pendulum which takes twelve hours. The information obtained is exactly the same as with the pendulum, the total intensity of the gravitational field. For this reason the presence of a heavy mass near by on the surface or at small depth has no effect on the reading; the presence of such a mass merely alters the direction of the

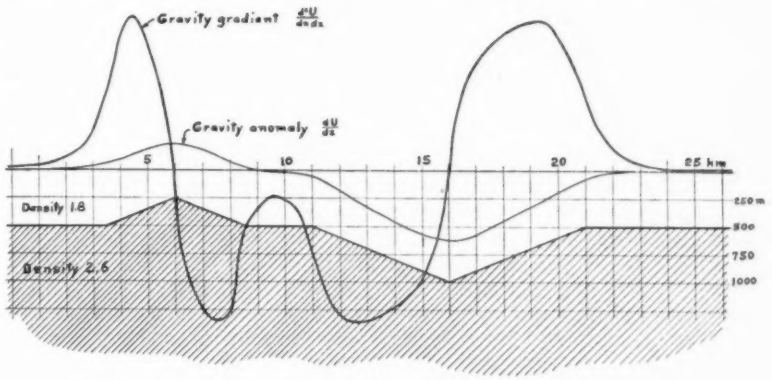

Fig. 4. Gradient and gravity curves for assumed structure.

resultant force and as the instrument is set exactly in line with this direction the measurement is the same. These surface masses make so much trouble for the torsion balance operator that it seems to be hard for him to believe that they do not disturb this instrument.

The relation between the gravity anomaly, which is measured directly by the gravity balance, and the gravity gradient, which is measured by the torsion balance, is very clearly shown in a diagram in Ambronn's Elements of Geophysics, which he has taken from a paper by von Eötvos. Figure 4 is derived from this diagram; the heavy curve representing the values of the gravity gradient and the light curve the gravity anomaly caused by a mass of heavy rock at varying depths below the surface. It will be noted that the curve representing the magnitude of the anomaly follows very closely the outline of the rock mass below while the curve representing the magnitude of the gradient bears no resemblance whatever to this outline and requires some study to interpret its meaning. In actual field work these curves are not obtained but only a rather small number of points on them. A comparatively

small number of gravity measurements would give the anomaly curve with a fair degree of accuracy but the same number of ordinates to the gradient curve would leave it still extremely uncertain and the results consequently very difficult to interpret and the interpretation of very doubtful validity.

While the new instrument is still in the experimental stage and there will undoubtedly be improvements in details of construction, it is not probable that it can ever be used effectively by an unskilled operator. The method of operation is extremely simple but it seems to be necessary for the operator to have had some kind of experience with precision measurement in order to guard against deceptive appearances and be sure that the instrument is really in equilibrium when the reading is made.

CHARTS FOR TORSION BALANCE READINGS*

By M. M. Slotnick
Geophysics Department, Humble Oil and Refining Company

(Received February 1, 1932)

THE calculations necessary to determine the values of the gradient and curvature quantities from the readings taken at a torsion balance station involve either slide-rule or logarithmic manipulations which are somewhat

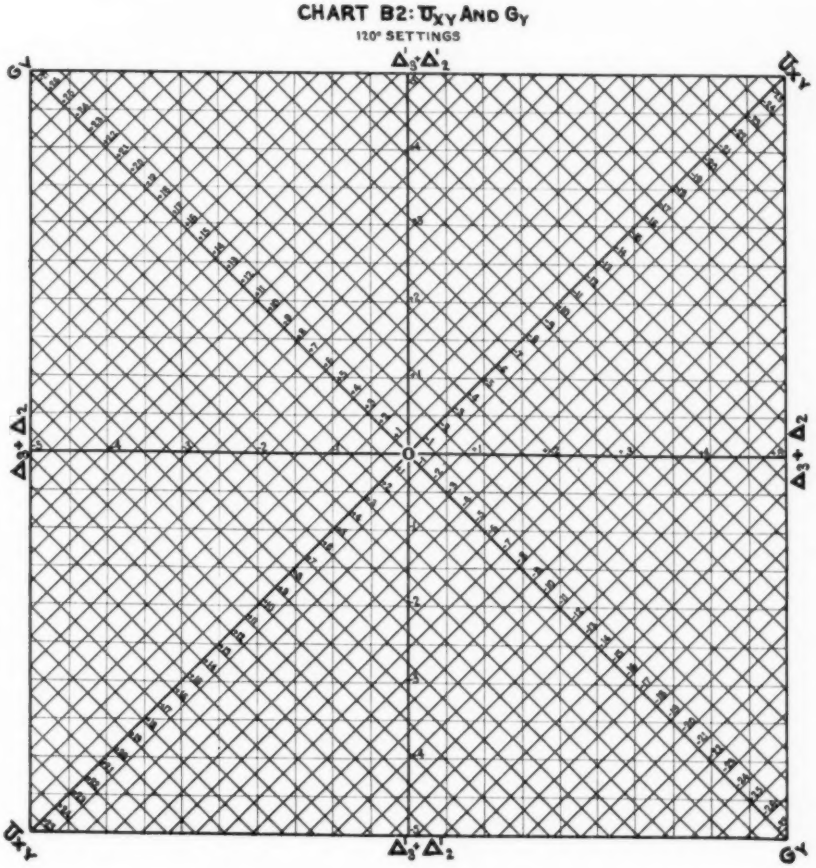

Fig. 1.

tedious, if not entirely cumbersome. It is quite a simple matter to show that the result obtained by writing the equations involved in linear forms of the type shown below is well within the accuracy of the instrument.

* Published by permission of *Humble Oil and Refining Company.*

A pair of charts can readily be made for each instrument when used in 72° azimuth settings and another pair for 120° azimuth settings. The charts appended were made for field use for a Bamberg instrument of the

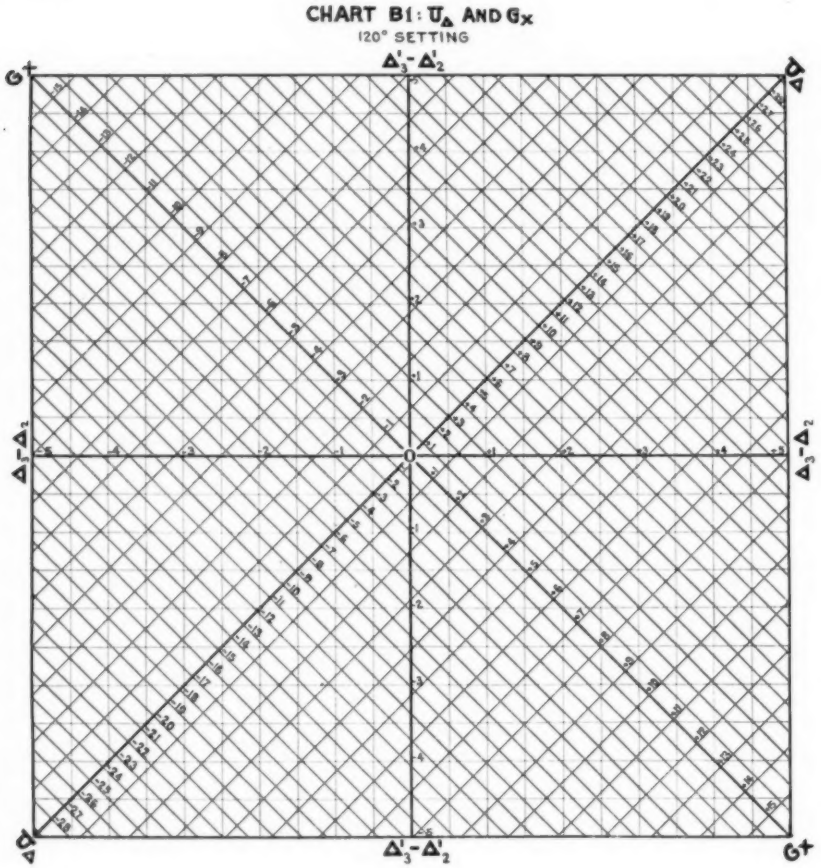

Fig. 2.

Humble Company equipped with Blau wires and are for the latter type of settings. For this particular instrument the equations in this case are:

$$G_x = 1.59[(\Delta_3' - \Delta_2') - (\Delta_3 - \Delta_2)],$$
$$U_\Delta = 2.93[(\Delta_3' - \Delta_2') + (\Delta_3 - \Delta_2)],$$
$$G_y = -2.75[(\Delta_3' + \Delta_2') - (\Delta_3 + \Delta_2)],$$
$$U_{xy} = -2.56[(\Delta_3' + \Delta_2') + (\Delta_3 + \Delta_2)].$$

The first pair of equations is charted on B1 and the second on B2. An examination of the charts will show at once the ease and accuracy with which they are read.

In accordance with the usual notation we have

$$\Delta_k = n_k - n_0, \quad k = 1, 2, 3,$$

where n_k is the reading of the instrument in its kth position, and

$$n_0 = \tfrac{1}{3}(n_1 + n_2 + n_3).$$

The unprimed letters refer to the quantities bearing on the one beam, and the primed to those on the other beam.

The originals of the charts are $20'' \times 20''$ with the horizontal and vertical scales running from -10 to $+10$ units at intervals of 0.1, a range well within that usually necessary.

It has been found that nomographs are not quite as convenient as these charts for field use.

THE EFFECTS OF HEAT TREATMENT ON FINE METALLIC SUSPENSIONS

By N. N. Zirbel
Department of Physics, The Rice Institute

(Received February 15, 1932)

Abstract

When a suspended system is supported by a fine wire the equilibrium position usually changes slowly for a long time after the load is applied. The equilibrium position also changes with temperature. It is found that both of these disturbing factors can be eliminated by a suitable heat treatment of the wire. Observations have been made on tungsten and platinum-iridium wires of sizes suitable for use in the Eötvos torsion balance.

Introduction

MANY types of physical apparatus involve the use of a rotating system suspended by a thin filament of metal or other material. Numerous irregularities in the behavior of an arrangement of this sort are often observed, such as changes in the equilibrium position with time and temperature, changes in the torsion constant of the suspending filament and failure of the suspended system to return to its initial position after a displacement. These irregularities are especially serious in the Eötvos torsion balance, in which the suspension is very thin and is heavily loaded. Investigations of various sorts on the behavior of suspension wires have been published by Shaw and Lancaster-Jones,[1] Boys,[2] Threlfall,[3] Königsberger[4] and others. The present work is devoted to an attempt to remove by suitable heat treatment the changes in equilibrium position with time and temperature. On this particular phase of the subject very little has been published.

Apparatus and Procedure

Observations have been made on several tungsten and platinum-iridium wires of diameters and lengths suitable for torsion balance use. The furnace used to anneal the wires and for the time drift and temperature coefficient determinations was a brass tube wound with an electrical heating coil and well insulated with asbestos. A glass window inserted in the side allowed the angular position of the suspended system to be observed with a lamp and scale, and a thermometer inserted at the top allowed the temperature to be read. In all the work the scale distance was 120 cm and the scale divisions millimeters.

[1] H. Shaw and E. Lancaster-Jones, Proc. Phys. Soc. London **35,** 151 (1923).
[2] V. C. Boys, Phil. Mag. **23,** 489 (1887).
[3] Richard Threlfall, Phil. Mag. **30,** 99 (1880).
[4] J. Königsberger, Zeits. f. Physik **40,** 729 (1927).

The suspended system used was a cylindrical brass weight with a mirror attached since this design would minimize the effects of convection currents. The weight of the suspended system was about half that required to break the wire. The upper end of the suspension was fastened to a brass rod which extended through the top of the furnace and was clamped outside so that a twisting or slipping of the furnace would not affect the position of the suspended system.

In the case of tungsten wires carbon dioxide gas was kept flowing through the furnace during all annealings at temperatures higher than 200°C. In order to anneal at temperatures higher than 500°C the wire was suspended in a glass tube with a weight of 6 to 10 grams at the lower end dipping in a copper sulfate solution so that a current could be passed through the wire.

The temperature coefficient, which is defined as the change in the equilibrium position per unit temperature change, was measured by heating the furnace about 20°C above room temperature and allowing it to cool slowly; the scale reading being noted at the beginning and at various temperatures during the cooling. Then a graph was plotted showing the scale reading as a function of the temperature. Since the relation between the scale reading and the temperature is linear, the temperature coefficient can be easily obtained from the graph. It was not possible to make readings during the heating because convection currents disturbed the suspended system in an irregular way and the alternating magnetic field of the heating coil exerted a torque on it. The possibility of convection currents affecting the position of the suspended system even during the slow cooling was investigated by making two sets of readings on the same wire, one set using the furnace in the usual manner, the other set by changing the temperature of the whole room very slowly. The corresponding readings were the same, within experimental error, showing that the effect of convection currents was not appreciable. For a typical temperature coefficient determination see Fig. 1. The coefficient is reckoned positive if an increasing temperature causes a clockwise rotation of the suspended system, as viewed from above, and negative if the rotation is counter-clockwise.

Experimental Results

The first observations were made at room temperature on changes in the equilibrium position with time. In one typical case the rate of drift immediately after loading the wire was 4.4 mm per hour. After twenty-four hours the motion was still easily observable but much slower, about 0.7 mm per hour. In some cases a change in the direction of the drift would occur. It seems that this time drift is caused by a slow release of strains of some sort in the wire and that in a sufficiently long time a stable condition would be reached. Just how long a time would be required is not known but Shaw and Lancaster-Jones[1] have given a case in which the drift was still going on after 50 days.

The next observations made at about 200°C showed that at this temperature the untwisting process was greatly accelerated so that the stable posi-

tion was reached in about 7 hours instead of many days. At higher temperatures still less time was required. During this baking process the total rotation was often 180° or more. It was also found that if a wire from which the time drift had been removed by baking were strained beyond the elastic limit by a longitudinal stress, then the time drift again appeared and could

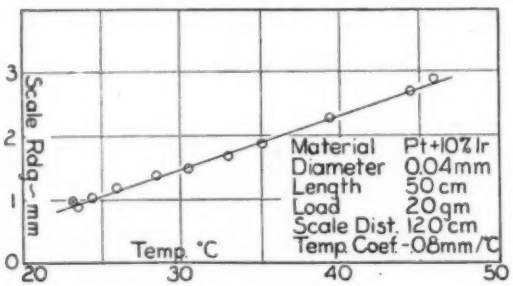

Fig. 1.

be removed only by further baking. The same result followed when the brass cylinder attached to the lower end of the wire was twisted through about half a revolution and held in this twisted position for several hours.

The time drift is thus quite easily removed. There remains, however, the change in equilibrium position with the temperature of the wire, as shown in Fig. 1.

The next problem is then to study the effect of heat treatment on this temperature coefficient, which in Fig. 1 is −0.08 mm per °C. Fig. 2 shows

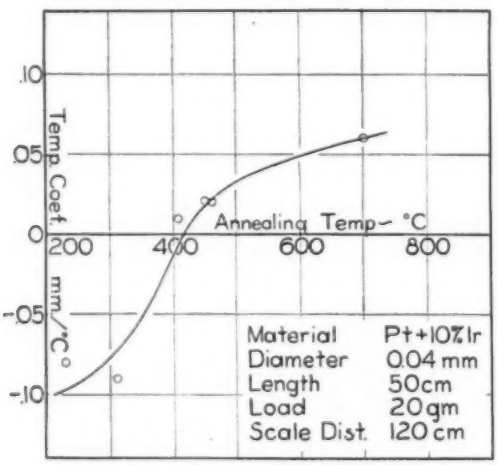

Fig. 2.

the results of such a study. The complete procedure in obtaining these results is as follows. The wire is hung up with a load of 20 grams and baked at 230°C for 6 hours. This baking removes the time drift. The temperature coefficient is then obtained by taking the readings shown in Fig. 1 and is found to be −0.08 mm per °C. This value gives the first point on Fig. 2. The

wire is then baked at 310°C for 2.5 hours. The temperature coefficient is then again measured between 20°C and 50°C and found to be −0.09 mm per °C. The process is then continued with successively higher baking temperatures. Fig. 2 shows that the temperature coefficient between 20° and 50°C becomes

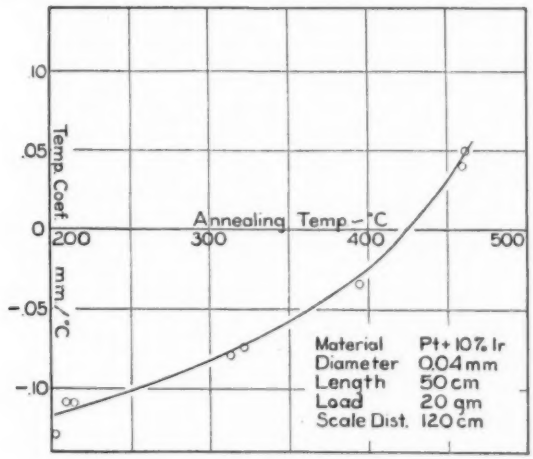

Fig. 3.

zero after baking at a certain critical temperature of about 425°C and takes on increasing positive values after baking at higher temperatures. Fig. 3 shows similar and somewhat more consistent results obtained with another length of the same wire. The critical temperature is about the same in the

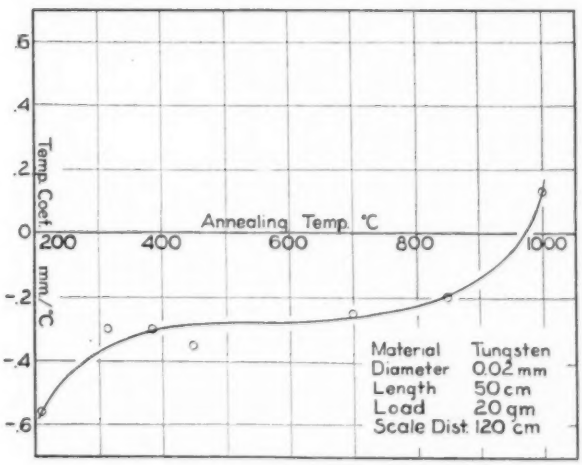

Fig. 4.

two cases. Fig. 4 shows that for tungsten the critical temperature is considerably higher. These latter results are subject to considerable error because temperatures higher than 500°C were estimated roughly by the color of the hot wire.

It has been found that the duration of time during which the wire is baked at a certain temperature is of little importance so far as the effect on the temperature coefficient between 20° and 50°C is concerned. In the results shown, this time varied from 3 minutes to 7 hours. It has also been found that after a wire has once been heated to a certain temperature, subsequent bakings at lower temperatures have no effect. That is, the curves shown can be traversed in one direction only, and after a wire has once been carried above the critical temperature no later heat treatment can diminish its temperature coefficient. Finally, measurements made on the two halves of the same wire show that the temperature coefficient is uniformly distributed along the wire and is not due to any local irregularity.

We have no explanation for the peculiar behavior shown in the last three figures. However, incomplete results on a fourth sample from an entirely different source show the same sort of behavior, and we believe that the effect, whatever its cause, is genuine. Suitable precautions have been taken to avoid the effect of convection currents and other disturbing factors. The fact that in all cases so far observed, involving wires from three sources, the temperature coefficient is initially in the same direction is also difficult to explain and may be merely a coincidence.

In conclusion, the writer wishes to express his thanks to Professor H. A. Wilson, under whose direction this work was done.

ON THE CORRELATION OF ISOGEOTHERMAL SURFACES WITH THE ROCK STRATA

By C. E. Van Orstrand[1]
United States Department of the Interior

(Received February 1, 1932)

Abstract

An instance of regional variation in Oklahoma and two cases of local variation, one at Long Beach, California, and the other at Salt Creek, Wyoming, have been selected for consideration from a large number of geothermal surveys conducted during the past few years by the U. S. Geological Survey and the American Petroleum Institute. The causes of local and regional variations are unknown. Possible explanations such as radioactivity, proximity to crystalline rocks, and transfer of heat along the strata are given careful consideration in attempting to explain the observed relations between the strata and the isogeothermal surfaces.

REGARDLESS of the source of the earth's internal heat, the mass of the earth is so large, and the time during which it has been losing heat is so long, that whatever the law of increase of temperature with depth may be, the theoretical temperature distribution to depths of one or two miles should be, to a very high degree of precision, a straight line; but, in a recent summary of the data from 400 deep wells, distributed in 18 states, only 5 percent of the depth-temperature curves could be classified as linear; 36 percent were concave to the depth axis; and 59 percent were convex to the depth axis. A curve of the latter type is shown in Fig. 1. The causes of these anomalies are not definitely known. Diminution of thermal conductivity with depth resulting from a corresponding diminution in the moisture content of the rocks may be a partial explanation of the observed convexity of the curves; and the increased conductivity of the more dense crystalline or basaltic rocks beneath thin sediments is a possible explanation of the concave curves. The curves tend to be more or less of the same type in each individual field. Not withstanding these anomalies, however, the curves obtained under ideal conditions are exceptionally smooth and uniform, and the resulting isothermal surfaces determined from them are likewise smooth and uniform.

Regional Variation of Isogeothermal Surfaces

The general trend of the 100°F isogeothermal surface[2] extending over a distance of 100 miles from Okemah to Oklahoma City, Oklahoma, is shown in Fig. 2. That these surfaces are closely related to the strata is evidenced by

[1] Published by permission of the Director of the U. S. Geological Survey.

[2] John A. McCutchin, Determination of geothermal gradients in Oklahoma. Bull. Am. Assoc. Petroleum Geologists **14,** 535–555 (1930); Determination of geothermal gradients in oil fields located on anticlinal structures in Oklahoma. Am. Petroleum Institute, Production Bull. No. 205, 19–29 (1930).

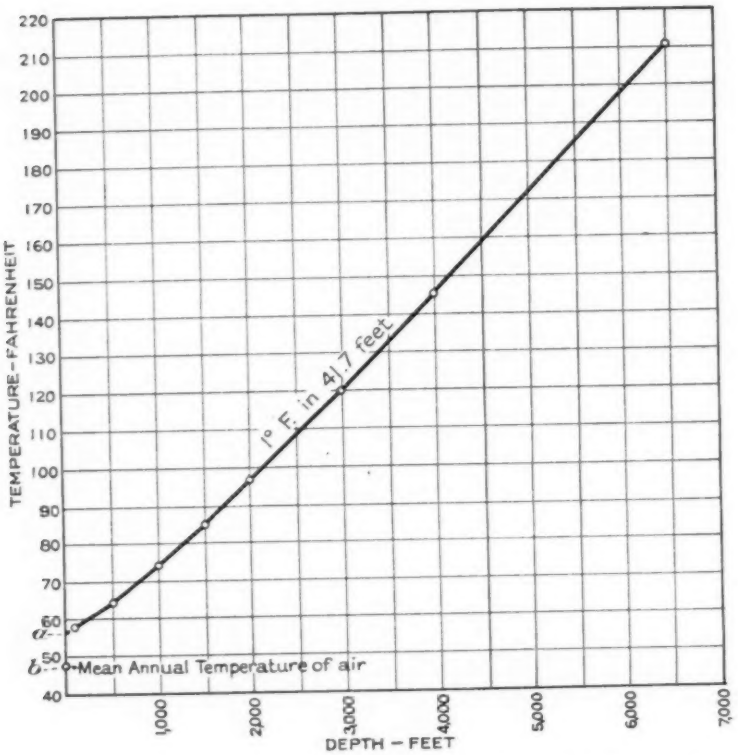

Fig. 1. Depth-temperature curve of deep well in S. W. 1/4, Sec. 34, T. 2 N., R. 69 W., Longmont, Boulder County, Colorado.

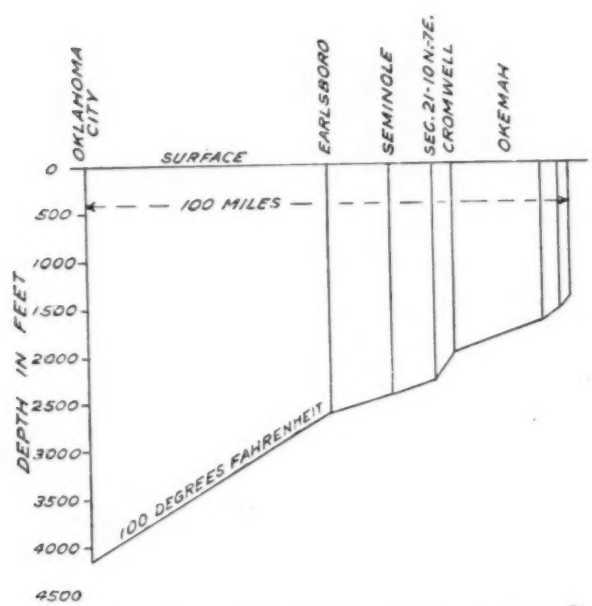

Fig. 2. Depths at which temperature of 100° Fahrenheit is reached in Central Oklahoma.

the fact that the crystalline rocks are found at a depth of about 2,500 feet at Okemah and at about 10,000 feet at Oklahoma City. The observations show that the isogeothermal surfaces tend to parallel the granitic bed at the base of the sedimentary strata. The irregularities which appear at several points along the general trend are apparently due to local uplifts on some of which oil fields are located.

Local Variations of Isogeothermal Surfaces

The Salt Creek and Long Beach oil domes, Figs. 3 and 4, are excellent examples of local uplift in which the isogeothermal surfaces stand in a definite

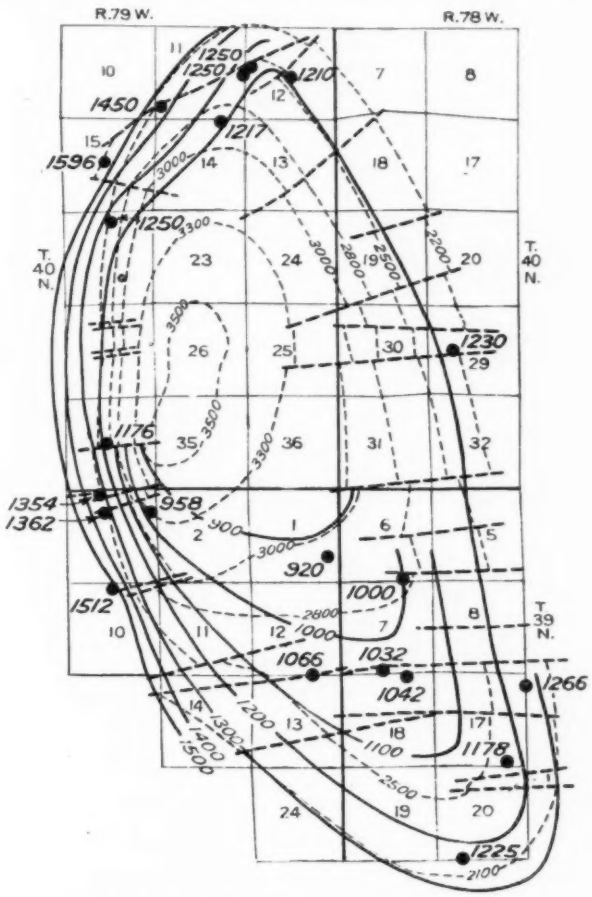

Fig. 3. Salt Creek Dome, Natrona County, Wyoming.

relation to the strata. In Fig. 3, the closed broken lines represent elevations on the "Second Wall Creek oil sand". Faults are represented by broken lines intersecting the contours. The small dots or disks represent well locations and the accompanying numbers represent the depths at which a temperature of 80°F is reached. The heavy continuous lines represent contours on the 80°F isogeothermal surface. Fig. 3 shows clearly that the highest point on the

isothermal surface corresponds very closely to the highest point on the "Second Wall Creek sand". Extending laterally from this point, the 100 and 120°F isothermal surfaces tend to parallel the surface of the sand as shown in

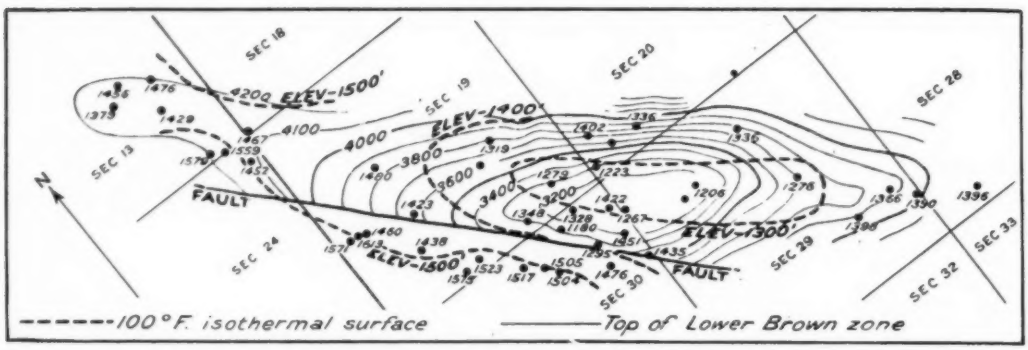

Fig. 4. Long Beach Dome, Los Angeles County, California.

Fig. 5.[3] In general, the isothermal surfaces are flatter than the surfaces of the strata and the flatness increases with diminishing depth until a point is reached near the surface of the ground at which the isothermal surfaces tend more or less to conform to surface topography. As the effects of surface

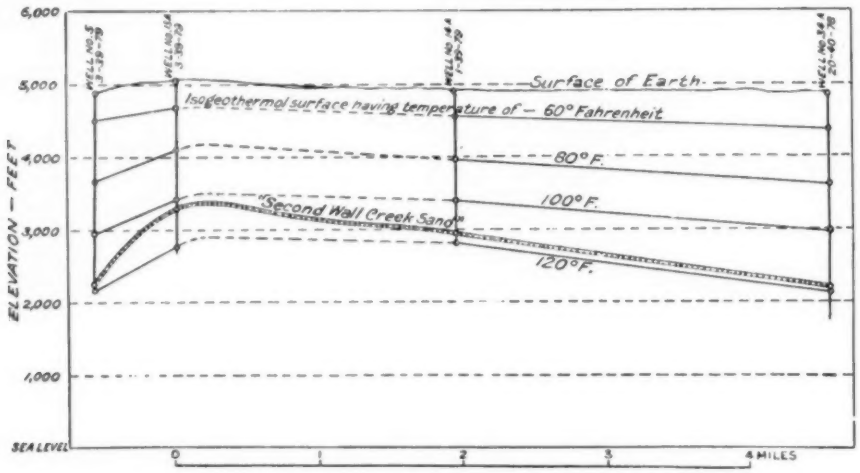

Fig. 5. Cross-section of Salt Creek Dome, Wyoming.

topography, however, are practically negligible in this field, we must look to other causes for an explanation of the marked rise of the thermal surfaces as they pass over the dome.

The Long Beach dome[4] presents a problem in which surface topography is

[3] W. T. Thom, Jr., Relation of earth temperatures to buried hills and anticlinal folds. Economic Geology **20,** 524–530 (1925).

[4] A. J. Carlson, Geothermal variations in oil fields of Los Angeles Basin, California. Bull. Am. Assoc. Petroleum Geologists **14,** 997–1011 (1930); Geothermal conditions in oil producing areas of California. Am. Petroleum Institute, Production Bull. No. 205, 109–139 (1930).

an important factor. As shown in Fig. 4, the isothermal surfaces are exceptionally smooth and uniform and follow the top of the upper oil producing

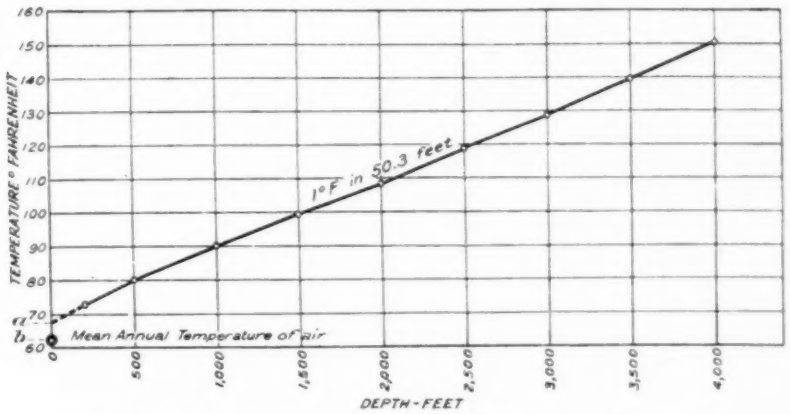

Fig. 6. Depth-temperature curve, deep well No. 2, Tucker and Johnson lease, Long Beach, California.

strata, the Lower Brown Zone, with marvelous precision. That uniformity in these surfaces was to be expected is suggested by the depth-temperature

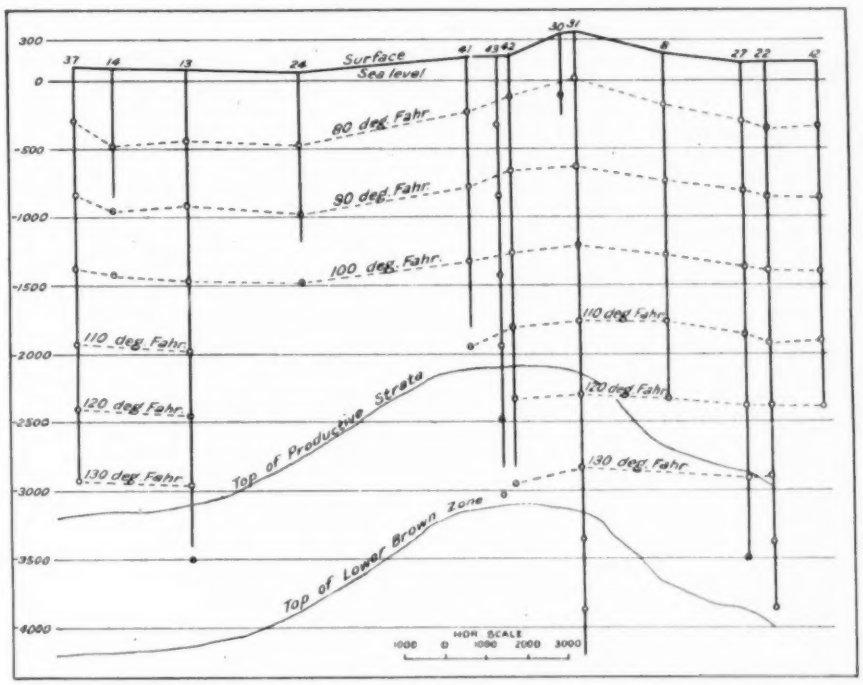

Fig. 7. Longitudinal section of Long Beach oil field. (See 4, first reference.)

curve of deep well No. 2, Fig. 6. This curve like many others in the Long Beach field, shows a slight rise in temperature at a depth of about 1000 feet,

beyond which the curve is slightly convex towards the depth axis. Another point of interest is the fact that the excess of annual mean soil temperature just beneath the surface of the ground over annual mean air temperature just above the surface of the ground (a–b, Fig. 6), amounts to more than 7°F. This large discontinuity in the depth temperature curves at the surface of the ground is typical of oil field areas in Southern California.

Correction for Surface Topography

The longitudinal section of the Long Beach field, Fig. 7, shows at once that the total rise in the isothermal surfaces as they pass over the dome and beneath the hill can not be explained on the basis of normal cooling; for, the

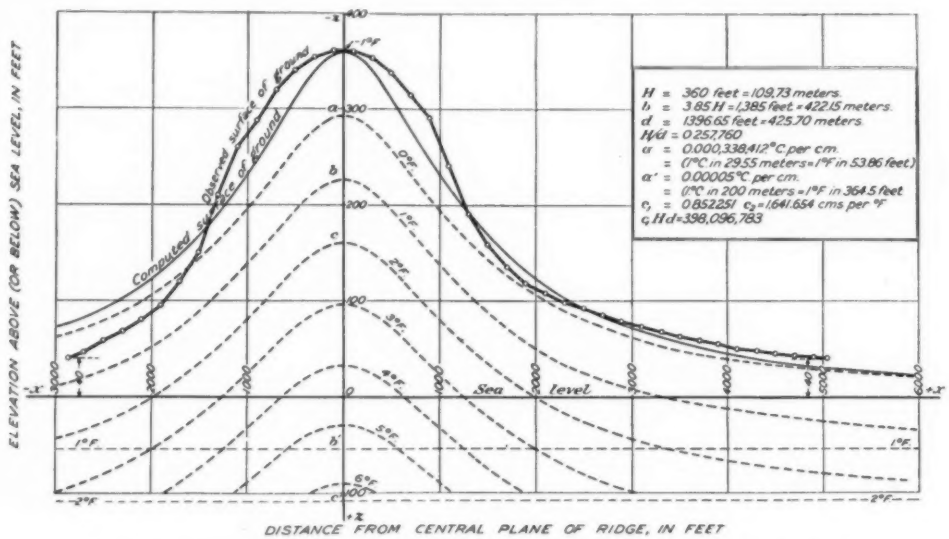
ELEVATION ABOVE (OR BELOW) SEA LEVEL, IN FEET
Observed surface of ground
Computed surface of ground
Sea level
DISTANCE FROM CENTRAL PLANE OF RIDGE, IN FEET
H = 360 feet = 109.73 meters.
h = 3.85 H = 1,385 feet = 422.15 meters.
d = 1396.65 feet = 425.70 meters.
H/d = 0.257760
α = 0.000,338,412°C. per cm.
= (1°C in 29.55 meters = 1°F in 53.86 feet.)
α' = 0.00005°C per cm.
= (1°C in 200 meters = 1°F in 364.5 feet.
c1 = 0.852251 c2 = 1641.654 cms. per °F
c1Hd = 398,096,783

Fig. 8. Transverse section of Long Beach Dome showing computed isotherms.

fact that the thermal surfaces are steeper and show a greater rise from the lowest to the highest points than the topographic surface above them proves conclusively that we are not dealing with a normal flow of heat towards the surface of the earth. In order to form an estimate of the discrepancy resulting from this abnormality, use will be made of the equations given by Lees.[5]

Referring to Fig. 8, let the axis of x be taken at sea level, with the origin o directly beneath the apex of the symmetrical hill. Elevations above sea level are negative; below, positive. As the hill or mountain is assumed to be a ridge which is very long in comparison with its cross-section, the problem is reduced at once to the flow of heat in two dimensions. The fundamental equation to be solved is

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial z^2} + \frac{w}{k} = 0 \qquad (1)$$

[5] Charles H. Lees, On the shapes of the isogeotherms under mountain ranges in radioactive districts. Proc. Roy. Soc. London **A83**, 339–346 (1910).

subject to the condition

$$k\frac{\partial v}{\partial n} = h(v - v_a) \tag{2}$$

which represents the loss of heat from the surface of the earth at any point on the plain or mountain slope.

In these equations, v = temperature at point x, z; v_a = temperature of air in contact with the surface of the ground at the given point; w = number of calories of heat generated per second per cubic centimeter by radioactive processes; k = coefficient of thermal conductivity of the rocks; h = coefficient of emissivity; n = numerical magnitude measured inward along the normal to the surface of the ground.

Now let us put, also, v_0 = annual mean temperature of soil just beneath the plain surface of the ground; α = temperature gradient beneath the plain; α' = temperature gradient in the air along the mountain slope.

$$\begin{aligned} c_1 &= (\alpha - \alpha')/\alpha \\ c_2 &= (v - v_0)/\alpha \end{aligned} \tag{3}$$

H = elevation of top of hill referred to level of plain; b = half-breadth of hill at elevation h nearly one-half the height of the hill; d = distance from apex of hill to point on contour at elevation h above the plain. (See Fig. 9).

$$\begin{aligned} \beta &= w/2k(\alpha - \alpha') \\ \frac{h}{H}\,&\frac{(1 + \beta h)}{(1 + \beta H)} = \tfrac{1}{2}. \end{aligned} \tag{4}$$

The quantity β is dependent on radioactivity. Eq. (4) is an arbitrary assumption introduced for convenience in handling the equations.

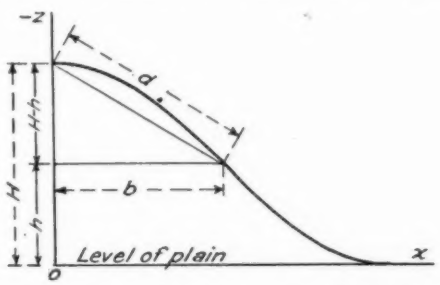

Fig. 9. Transverse section of imaginary hill or mountain.

Without going through the necessary steps of integration of Eqs. (1) and (2) and the subsequent transformation of the integrals in order to take into account the cross-section of the mountain, we write at once the equation to the contour of the transverse section,

$$z(1 - \beta z) + Hd(1 + \beta H)\frac{(z + H + d)}{x^2 + (z + H + d)^2} = 0 \tag{5}$$

and the corresponding equation to the isogeotherms,

$$v - v_0 = \alpha z - \frac{wz^2}{2k} + (\alpha - \alpha') \frac{Hd(1 + \beta H)(z + H + d)}{x^2 + (z + H + d)^2} . \quad (6)$$

As Eqs. (5) and (6) are cubics in z, it will be more convenient for purposes of calculation to put them in the respective forms:

$$x^2 = - [z + H + d] \left[(z + H + d) + \frac{Hd}{z} \right] \quad (7)$$

$$x^2 = - [z + H + d] \left[(z + H + d) + \frac{c_1 Hd}{z - c_2} \right] \quad (8)$$

in which the terms dependent on radioactivity have been neglected.

Putting $x = 0$ in (8),

$$z^2 + (H + d - c_2)z - (H + d)c_2 + c_1 Hd = 0 \quad (9)$$

and assigning particular values to c_2, which according to Eq. (3) represents the depth beneath the plain at which a particular temperature v is reached, Eq.

TABLE I. *Depth of isogeotherms beneath level of plain and rise of isogeotherms beneath apex of hill.*

Isogeotherms		Depth of isogeotherms beneath level of plain c_2		Rise of isogeotherms beneath apex of hill		Isogeotherms Degrees		Depth of isogeotherms beneath level of plain c_2		Rise of isogeotherms beneath apex of hill	
° C	° F	Meters	Feet	Meters	Feet	Cent.	Fahr.	Meters	Feet	Meters	Feet
0.0	0	0.0	0	89.2	293	38.9	70	1149.2	3770	24.0	79
1.1	2	32.8	108	81.8	269	41.7	75	1231.2	4039	22.8	75
2.2	4	65.7	215	75.8	249	44.4	80	1313.3	4309	21.8	71
3.3	6	98.5	323	70.7	232	47.2	85	1395.4	4578	20.9	68
4.4	8	131.3	431	66.3	218	50.0	90	1477.5	4847	20.0	66
5.6	10	164.2	539	62.5	205	55.6	100	1641.6	5386	18.5	61
6.7	12	197.0	646	59.1	194	61.1	110	1805.8	5925	17.1	56
7.8	14	229.8	754	56.1	184	66.7	120	1970.0	6463	16.0	52
8.9	16	262.7	862	53.5	175	72.2	130	2134.2	7002	15.0	49
10.0	18	295.5	969	51.1	167	77.8	140	2298.3	7540	14.1	46
11.1	20	328.3	1077	48.9	160	83.3	150	2462.5	8079	13.3	44
13.9	25	410.4	1346	44.1	145	88.9	160	2626.6	8618	12.6	41
16.7	30	492.5	1616	40.3	132	94.4	170	2790.8	9156	12.0	39
19.4	35	574.6	1885	37.1	122	100.0	180	2955.0	9695	11.4	38
22.2	40	656.7	2154	34.4	113	105.6	190	3119.1	10233	10.9	36
25.0	45	738.7	2424	32.1	105	111.1	200	3283.3	10772	10.5	34
27.8	50	820.8	2693	30.0	98	116.7	210	3447.5	11311	10.0	33
30.6	55	902.9	2962	28.2	93	122.2	220	3611.6	11849	9.6	32
33.3	60	985.0	3232	26.7	87	127.8	230	3775.8	12388	9.3	30
36.1	65	1067.1	3501	25.2	83	133.3	240	3940.0	12926	8.9	29

$\alpha = 0.000\ 338\ 412$°C per cm (=1°C in 29.55 meters=1°F in 53.86 feet).
$\alpha' = 0.000\ 05$°C per cm (=1°C in 200 meters=1°F in 364.5 feet).[6]
H=360 feet=109.73 meters.
$b = 3.85H = 1385$ feet=422.15 meters.
$d = 1396.65$ feet=425.70 meters. $H/d = 0.257\ 760$.
$c_1 = 0.852\ 251$ $c_2 = 1641.654$ cm per degree Fahrenheit.

[6] W. J. Humphreys, Physics of the Air. Sec. ed., p. 40 (1929).

(9) gives the elevation beneath the apex of the hill of that particular isotherm which is asymptotic to c_2. The difference between z and c_2 is equal to the rise of the isothermal surface above its asymptotic plane. Thus in Fig. 8, *oa* represents the rise of that isothermal surface which is just beneath the surface of the plain, ($c_3=0$), and which has *ox* extended for an asymptote; *bb'* represents the rise of that surface whose asymptote is perpendicular to the axis of the hill at *b'*; and so on. These values are tabulated in Table I. The table applies to any other hill of the same cross-section and the same geothermal gradients (α and α') beneath the plain and along the mountain slope. To determine the temperature represented by any isotherm, it is merely necessary to add the tabular value in the table, or Fig. 8, to the annual mean temperature of the ground just beneath the surface of the plain.

In order to compare the theoretical rise of the isogeothermal surfaces with the observed, the gradient beneath the plain was assumed to be, $\alpha = 0.000{,}338{,}412$, corresponding to a rate of temperature increase of 1°C in 29.55 meters, or 1°F in 53.86 feet. This value is an approximate average of the gradients for 42 wells in the Long Beach field. From the same wells, the average annual mean temperature just beneath the surface of the ground was found to be, $v_0 = 71.6 \pm 0.2$°F. Hence, the 80°F isotherm corresponds to a temperature difference in Table I of 8.4°F. Interpolating, we find the rise of the isotherm to be 215 feet. The last two columns of Table II were obtained by a repetition of this process, using for the last column a table similar to Table I. In column 4, the level of the plain was assumed to coincide with sea level; in the last column of the table, it is taken to be 40 feet above sea level. The data for the transverse section of the hill, Fig. 8, were kindly forwarded to me by Dr. A. J. Carlson. The observed differences in elevation of the surface of the ground and the isotherms, column 3, Table II, are taken from his paper on "Geothermal variation in the oil fields of Los Angeles Basin, California."[4]

Comparison of column 3 of Table II with the last two columns shows that the observed heights are somewhat more than twice the computed

TABLE II. *Rise of isogeotherms.*

Isogeotherms °F	Depth beneath plain feet	Differences in elevation		
		Observed feet 296	Assumed Computed feet 360	Assumed Computed feet 320
80	453	492	215	189
90	991	348	166	144
100	1530	274	136	118
110	2068	223	116	100
120	2607	182	100	86
130	3146	200	89	76
140	3684	141	80	68
150	4223	—	72	62
Lower Brown Zone	—	1100	—	—

heights. In making this comparison, I have attempted not to underestimate the computed heights. For example, the assumed elevations of the top of the hill above the level of the plain, namely 360 and 320 feet, instead of 296 feet given by Dr. Carlson, gives a rise of isothermal surface which is too high. Likewise, the substitution of a long ridge for a hill tends to make the computed values too high. An important possible exception, however, should be noted. Let us assume that the gradient beneath the plain is the steepest in the field corresponding to a rate of 1°F in 48.5 feet instead 1°F in 53.86 feet as used in constructing Table I. Hence we have

$$\alpha = 0.000\ 375\ 811 \quad c_1 = 0.866\ 954$$

and substituting in Eq. (9), putting $c_2 = 0$ in order to represent the 0°F isotherm, we find 299 feet instead of 293 feet as recorded in Table I. This difference of 6 feet is rather insignificant in comparison with the large difference of more than 100 feet for which we require an explanation. Using the value $\alpha' = 0.00006$ instead of $\alpha' = 0.00005$, the corresponding rise is 288 feet instead of 299 feet. This shows that the value of α' is of importance in determining the computed heights. Taken as a whole, however, the evidence substantiates the conclusion that the differences in elevation between the highest and lowest points on the observed isotherms are abnormally large.

Comments on the Data of Observation

The tests in the Salt Creek field were made under conditions such that a close approximation to true rock temperatures is to be expected. Values of a in the equation

$$y = a + bx$$

in which y = temperature at depth x; a = annual mean temperature just beneath the surface of the ground; b = gradient, vary from 46.9 to 52.3°F. The mean from 19 wells is 49.8 ± 0.2°F. The probable error r of an observation of weight unity is 1.0°F. Values of the excess of soil temperature over air temperature ($a - b$, Figs. 1 and 6) range from 0.2 to 5.6°F. The mean is 3.10 ± 0.2°F. $r = \pm 1.0$°F.

At Long Beach, the temperatures just beneath the surface of the ground from 42 wells fall between the limits 68.4 and 77.2 with a mean value of 71.6 ± 0.2°F. $r = \pm 1.4$°F. The maximum value of the excess of soil temperature over air temperature is 14.6, the minimum, 5.8°F. The mean is 9.0 ± 0.2°F. $r = \pm 1.4$°F.

The large range in the preceding values suggests the possibility that the temperatures in some of the wells that had been pumped are too high. The resulting error is a maximum at or near the surface of the ground and gradually diminishes with more or less regularity to zero at the source of the rising column of oil or water. An obvious effect of this abnormality is to elevate the isotherms in amount roughly proportional to their elevation above the fluid bearing sand, and diminish the gradients or increase the reciprocal gradients in the same relative proportion. To ascertain the possible magnitude of this

TABLE III. *Mean values of a, b, and 1/b.*

100–1000 feet								100–2000 feet				
No. of Obs.	a	r	r_0	b	$r \times 10^5$	$r_0 \times 10^5$	$1/b$	No. of Obs.	b	$r \times 10^5$	$r_0 \times 10^5$	$1/b$
5	72.8	1.2	0.5	0.01781	58	26	56.1	5	0.01776	44	20	56.3
35	71.3	1.4	0.2	.01885	9	2	53.0	35	.01878	59	10	53.2
10	73.4	1.6	0.5	0.01804	87	28	55.4	10	0.01804	63	20	55.4
30	70.9	1.0	0.2	.01895	88	16	52.8	30	.01885	55	10	53.0
10	73.4	1.6	0.5	0.01804	87	28	55.4	10	0.01804	63	20	55.4
10	70.9	1.0	0.3	.01919	81	26	52.1	9	.01923	34	11	52.0
10	71.2	1.1	0.3	.01855	108	34	53.9	10	.01849	69	22	54.1
10	70.4	1.1	0.3	.01912	74	23	52.3	9	.01882	38	13	53.1

source of error, Table III has been computed. The table is based on the data from 40 wells, two of the poorest values being omitted. In the first two lines, the mean values for 5 wells highest on the structure are compared with the means from the remaining 35 wells. Similarly, in the third and fourth lines, the means for 10 wells highest on the structure are compared with the means from the remaining 30 wells, while in the final tabulation the wells have been grouped in four successive groups as regards elevation on structure, the first

TABLE IV. *Gradients and reciprocal gradients (H = 360 feet; and H = 320 feet).*

In air on slope of hill			Beneath adjacent plain			Beneath summit of hill		
Gradients (α')	Reciprocal gradients		Gradients (α)	Reciprocal gradients		Gradients (α_s)	Reciprocal gradients	
°C per cm	Feet per °F	Meters per °C	°C per cm	Feet per °F	Meters per °C	°C per cm	Feet per °F	Meters per °C
0.00006	303.8	166.7	0.00036	50.6	27.8	0.00028 27	64.5	35.4
0.00006	303.8	166.7	0.00030	60.8	33.3	0.00023 81	76.5	42.0
0.00005	364.5	200.0	0.00036	50.6	27.8	0.00028 01	65.1	35.7
0.00005	364.5	200.0	0.00030	60.8	33.3	0.00023 56	77.4	42.5
0.00006	303.8	166.7	0.00036	50.6	27.8	0.00028 64	63.6	34.9
0.00006	303.8	166.7	0.00030	60.8	33.3	0.00024 11	75.6	41.5
0.00005	364.5	200.0	0.00036	50.6	27.8	0.00028 40	64.2	35.2
0.00005	364.5	200.0	0.00030	60.8	33.3	0.00023 87	76.4	41.9

group as heretofore being the highest. The table shows clearly that the values of a and $1/b$ in the first group of each set are always the highest. However, minor irregularities of this kind can not always be attributed to pumping operations. In general, the depth-temperature curves become more irregular as the surface of the ground is approached. It is for this reason that a map showing the isothermal surfaces is much more reliable than a map showing the distribution of reciprocal gradients. In addition to the preceding irregularities, the variation of the annual mean temperature of the air over the

surface of the hill must be taken into account. Thus, the value of a for six wells on the south side of the hill is 73.0 ± 0.8; the same for 15 wells on the north side is 71.4 ± 0.3, making a possible temperature difference on the two slopes of 1.6 ± 0.9°F. Other observers have likewise found maximum temperatures on southern slopes.[7]

As a further explanation of the slight increment in the observed values of $(1/b)$ as the crest of the hill is approached, a comparison must be made between the assumed gradient (α) beneath the plain and the theoretical gradient (α_s) beneath the summit of the hill. Using the value of H/d given in Table I and Fig. 8, and making the appropriate substitution in the equation

$$\alpha_s = \alpha - H(\alpha - \alpha')/d$$

we obtain the values of α_s tabulated in the first part of Table IV. In the second part of the table, (α_s) is given for $H = 320$ feet, $H/d = 0.245\ 240$.

Interpolating for our assumed reciprocal gradient beneath the plain, 1°F in 53.86 feet, we find the value 1°F in 68.3 feet at the summit of the hill. That is, the computed variation from the plain to the summit of the hill is four or five times the differences, 3.1 and 2.3, shown between the highest and lowest groups on the structure in the last set of values in Table III. In one way only does it seem possible to bring theory and observation into agreement. Substitution in the equation,

$$\frac{\partial v}{\partial z} = \alpha - (\alpha - \alpha') \frac{Hd}{(z + H + d)^2} \tag{10}$$

which represents the gradient at any elevation (z) beneath the apex of the hill, we find for

$$\alpha = 0.00033\ 8412 (1° \text{ F in } 53.86 \text{ feet}),\quad \alpha' = 0.00005,\quad H = 360 \text{ feet},$$

the values,

$1/b = 1$°F in 62.5 feet	$z = 0$ meters
$= $ " 55.9 "	$z = 500$ "
$= $ " 54.8 "	$z = 1000$ "

At $z = o$, 360 feet beneath the top of the hill, the theoretical difference of 8.6 feet between the assumed and computed values of $(1/b)$ greatly exceeds the observed differences, 3.1 and 2.3 feet, Table III; but at $z = 500$ meters, corresponding to a depth of 2000 feet beneath the top of the hill, the observed differences slightly exceed the computed value of 2.0 feet. At greater depths, Eq. (10) shows that the gradients beneath the plain and the hill approach coincidence.

To sum up the evidence at Long Beach, the elevation of the isotherms with reference to the topographic surface, and the small variation of gradient over the structure in comparison with the theoretical variation proves con-

[7] Edith M. Fitton, Soil temperatures in the United States. Monthly Weather Review **59**, 9 (1931).

clusively that the temperature distribution in this field is highly abnormal. The only exception appears to be the possible agreement of the observed and computed gradients at depths of about 2000 feet.

At Salt Creek, the evidence of variation of temperature with structure is conclusive. At Long Beach, the evidence is confused on account of the fact that in some of the wells the temperatures at the higher levels were increased by pumping operations while in a few of the remaining wells, the temperatures may have been lowered as a result of the wells being on a vacuum. However, proceeding on the assumption that we have a normal distribution of temperatures, and taking into account the fact that the abnormal temperatures in some of the wells have tended to increase the values of $(1/b)$, it follows that the observed variation in the values of $(1/b)$ as we pass from the plain to the summit of the hill, should have exceeded slightly the theoretical variation; but, just the converse is true—the observed variation is only a small fraction of the theoretical variation. Hence we are driven to the conclusion that abnormal temperature conditions in this field have tended to neutralize the effects of surface topography.

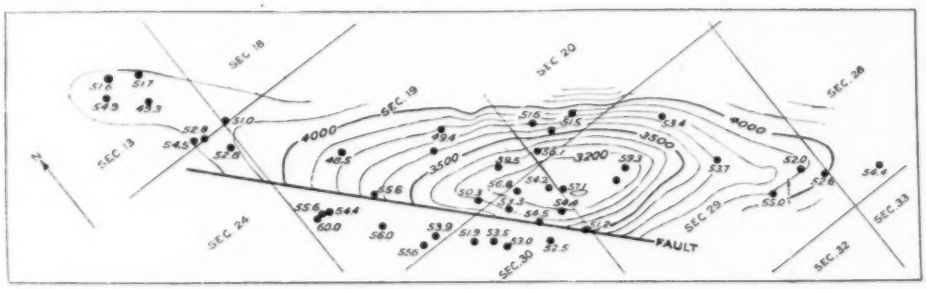

Fig. 10. Long Beach Dome. The numbers opposite well locations represent feet per degree Fahrenheit, 100–2000 feet.

Causes of Temperature Variations

Depth to the crystalline rocks and flow of heat along the strata appear to be possible explanations of the close relationship that apparently exists between the isotherms and the general trend of the strata in central Oklahoma (See Fig. 2).

A similar explanation may apply at Salt Creek, which is an exceptionally well defined local uplift. Steep dips occur on the western side of the field (See Figs. 3 and 5) and very hard granite was found by The Midwest Refining Company at a depth of 5420 feet near the top of the dome in the S.W. 1/4 N.W. 1/4, sec. 35, T. 40 N., R. 79 W. This evidence suggests the possibility that the depth to the granite is least on top of the dome, and that the direct contact of the granite with the hot crystalline rocks at great depths serves to maintain lines of flow from an almost inexhaustible heat source to the top of the dome. Additional evidence in substantiation of this conclusion is to be found in the fact that the values of $1/b$ vary from 1°F in about 29 feet on top of the dome to 1°F in about 56 feet in Sec. 15, T. 40 N., R. 79

W. on the outer edge of the dome. Not only is the great range in values of $1/b$ in agreement with our hypothesis, but the further fact that abnormally high temperatures are found immediately above the granitic mass leaves little to be desired in explanation of the observed distribution of temperatures.

On account of the great depth to the crystalline rocks at Long Beach, possibly 20,000 feet, it seems unlikely that depth to the crystalline rocks, which appears to be an important factor at Salt Creek and in Central Oklahoma, will provide an explanation of the temperature variations found in this field. Steeply dipping strata are present, but about the only other possibilities to which we can appeal are recent erosion, abnormal thermal conductivities and generation of heat in the strata. I shall not here attempt to discuss the merits of these possibilities.

To estimate the effects of radioactivity on the isogeotherms at Long Beach, let us expand Eq. (4) into the form

$$h = \frac{H}{2}\left[1 + \frac{\beta}{2}H + \cdots\right]$$

and then putting

$$H = 360 \text{ feet} = 10973 \text{ cms} \quad w = 1.5 \times 10^{-13}$$
$$\alpha = 0.00036 \quad \alpha' = 0.00006 \quad k = 0.004$$

we have

$$h = \frac{H}{2} = 5486 \text{ cms} \quad \beta = \frac{w}{2k(\alpha - \alpha')} = 6.25 \times 10^{-8}$$

and the new value of h is found to be 5488 cms an increase of only 2 cms. Using Joly's[8] recent value, $w = 42.1 \times 10^{-14}$, we have

$$\beta = 1.75 \times 10^{-7} \quad h = 5491 \text{ cms}$$

an increase of 5 cms. The resulting change in d (Fig. 9) is obviously negligible, consequently Eq. (6), which contains (d), remains unchanged.

Another method of procedure consists in computing the rise in the 0°F isotherm beneath the apex of the hill. Putting $x = o$ and $z = o$ in (6) and using the same constants as before, including Joly's value of w, we find

$$v + \Delta v = (\alpha - \alpha')(1 + \beta H)\frac{H}{1 + \frac{H}{d}}$$
$$= 2.617 + 0.005 = 2.622° \text{ C}$$

from which it again appears that the term due to radioactivity is negligible in comparison with the observed variations in temperature across the structure.

[8] John Joly. The surface-history of the earth. Second edition, p. 60 (1930). The Clarendon Press, Oxford.

Some General Suggestions

Our observations are not yet sufficiently accurate nor sufficient in number to permit making definite generalizations. The chief source of inaccuracy is unstable temperature equilibrium in wells. Notwithstanding these irregularities, however, the observations show conclusively that the isogeothermal surfaces rise in passing over a large number of salt domes and anticlinal structures. There is no well established evidence of depressed isogeotherms in any oil field, but in a considerable number of fields, the rise in temperature is not very marked. In general, the observations suggest the conclusion that a relation exists between the high temperatures and positive gravity anomalies that prevail over the crests of many of the oil field anticlines.

The causes of temperature variations given in the preceding paragraphs are to be regarded merely as preliminary suggestions. The complete explanation will undoubtedly involve isostasy, geological history, and the sources of the earth's internal heat.

GEOTHERMAL GRADIENT DETERMINATIONS IN THE LAKE SUPERIOR COPPER MINES

By L. R. Ingersoll
University of Wisconsin

(Received February 10, 1932)

Abstract

The Michigan College of Mining and Technology, in cooperation with the Calumet and Hecla Copper Company and the author, is carrying out a program of temperature measurements in the deep copper mines of Northern Michigan, extending the previous work of Agassiz and others. Temperatures are measured with mercury thermometers mounted in bakelite tubes, placed in drill holes in mine workings where the rock has been freshly exposed, special attention being given the effects of drilling, blasting, and other heat conduction considerations. Present results give as the average gradient from the surface to 5679 feet below (temp. 95.3°F), 1°F in 108.5 feet (0.0168° C/meter). The gradient is more nearly uniform than has sometimes been supposed. A preliminary attempt has been made at calculating the previous "thermal history" of this region. Diffusivity of specimens of the rock measures 0.0075 c.g.s., and on this basis calculations of theoretical temperature-depth curves have been made for 25 different assumptions of previous temperature conditions, and compared with the actual curve. Results as yet are inconclusive but indicate that at least 30,000 years have elapsed since the last glacial epoch, a longer period than usually assumed.

EVER since the days of Kelvin's theoretical discussion of the age of the earth there has been interest in measurements of the earth's temperature gradient, and while radioactive discoveries have considerably diminished the importance of calculations on this basis there are nevertheless many other interests—some of them new—attaching to geothermal measurements. As a result a mass of thermal data has been and is being collected from deep mine, well and drill hole measurements in all parts of the world. The copper mines of the Keweenaw peninsula in northern Michigan offer a particularly inviting opportunity for such study because of their great depth and freedom from local heat sources. The pioneer measurements of Agassiz,[1] which resulted in an apparently unusually low gradient, are too well known to require comment, and Lane,[2] Van Orstrand[3] and others have also made valuable contributions along this line.

Certain favorable opportunities, depending partly on the increasing depth of the mines, have for some time called for a reopening of this study and accordingly Professor James Fisher of the Michigan College of Mining and Technology and Mr. Harry Vivian, Chief Engineer of the Calumet and Hecla

[1] Alexander Agassiz, Am. Jour. Sci. (3) **50,** 503 (1895): Brit. Assoc. Adv. Sci. Report of 71st meeting, p. 65 (1901).

[2] A. C. Lane, Bull. Geol. Soc. of Am. **34,** 703 (1923).

[3] C. E. Van Orstrand, Am. Jour. Sci. **15,** 495 (1928). See also N. H. Darton, U. S. Geol. Survey, Bulletin 701, p. 50.

Consolidated Copper Company, are cooperating with the writer in a program of work which may extend over a period of years. The Michigan College is generously supplying the financial aid and the Calumet and Hecla Company the skilled technical assistance necessary to the success of the work. The measurements in all cases are to be guided by heat conduction considerations to the end that they may represent as closely as possible the actual virgin temperature at the spot, unaffected by mining operations. The work must accordingly proceed slowly, waiting frequently for favorable opportunities for temperature measurement; but while it has been going only about a year some preliminary measurements of value have been obtained as well as other results which may be of interest to the physicist.

Method of Temperature Measurement

General procedure

The measurements so far have been confined to the Calumet and Hecla mine in which most of the Agassiz measurements were also made. A somewhat different procedure has been adopted, however. In the Agassiz work thermometers were sealed for weeks at the bottom of drill holes 10 feet deep, in shafts or passages which had, themselves, been exposed to ventilation in some cases for many months. A little calculation on the basis of heat conduction theory will serve to show that the results under such conditions are very likely to be influenced by the ventilation, and only in the most favorable cases can be depended on to give virgin rock temperatures.

In the present work the measurements have all been carried out in new workings, i.e., drifts or other cuttings which have been advancing steadily a number of feet a week and which are well removed from other parts of the mine. At such a "temperature station" a hole is drilled a few feet back from the breast and in rock whose face has been exposed only a few days. The hole is usually 7 feet deep, but in some cases special 14 foot holes are run. Temperatures are taken with two or more thermometers located at the bottom of the hole. It may be remarked that, save for the water used in drilling, which quickly drains out, the hole has been found practically dry in all cases.

Thermometers

After careful consideration of the advantages and disadvantages of electrical temperature measuring instruments, mercury thermometers have been finally chosen for this work, at least for the type of measurements being made at present. They have been specially made by Henry J. Green, reading $0-40°C$ in $1/10°$, and frequent zero tests and comparisons with one of the number calibrated by the Bureau of Standards give maximum errors of the order of $0.04°C$. They are mounted in bakelite tubes 1 inch in diameter and 14 inches long with the thermometer bulb specially insulated thermally. This is a vital matter to which a great deal of attention has been given. It is necessary that after pulling from a drill hole the mercury thread show no change for something like a minute, so as to give the operator plenty of time to make his readings. At the same time there must be combined with this lag or "delay

action" a reasonably rapid action of the thermometer, once the bulb starts to change in temperature, as otherwise it would occupy too much time in coming to the temperature of the hole.

After some calculation and much experimentation a very satisfactory solution of the problem has been reached, involving insulating the bulb in vulcanite, wax and dynamo paper with waterproof lacquer as shown in Fig. 1. If such a thermometer is subjected to even such a radical temperature change as is involved in plunging into ice water, it is a full minute before the mercury thread shows a change of 1/10°. Having started to move, however, it falls several degrees a minute and comes down (asymptotically) to the

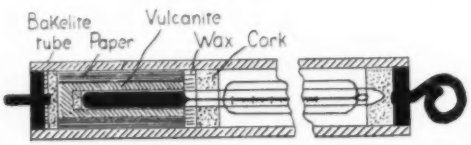

Fig. 1. Section of mounted thermometer (Overall length, 35 cm).

temperature of the bath—within 0.02°, say—in about half an hour. The final design of these thermometers has proved very satisfactory indeed for the work in hand. Difficulties experienced with the early ones over the separation of the mercury thread have been entirely overcome in the later ones by having them made with about one atmosphere of inert gas pressure over the mercury. This has the further advantage of relieving the bulb of practically all pressure, inside or outside, and doubtless tends to render the zero still more stable.

Temperature readings

As the holes are drilled in the usual manner with air-operated machines, with a flow of water through the drill steel, there is almost no heating effect due to drilling. Temperature readings taken immediately after the hole is completed are sometimes a degree high, but after a few hours all evidence of drilling heat has usually disappeared. To be on the safe side, however, little reliance is placed on readings made less than 24 hours after drilling, and the whole question of possible small errors due to heat of drilling is one on which considerably more study is to be expended.

Thermometers are used in tandem groups of two or three, run into the hole with a stout wire, with the last one carrying a rubber gasket. The mouth of the hole is stopped with a wooden plug. While the thermometers reach equilibrium in half an hour they are left for at least an hour and then quickly withdrawn and read. They are then replaced in the hole, to be read again about every two hours until several sets have been taken. This procedure is repeated a couple of days later, with occasional readings continuing for several weeks or months. In addition, air, rock face and psychrometer readings are occasionally taken in the drift to furnish general information as to the thermal conditions at the station.

Plots of these temperatures with time usually give an approximately

straight line, sloping slowly downwards, showing the gradual cooling—of the order of a degree or two a month—which takes place as heat is conducted from the rock into the slightly cooler drift. A slight extrapolation of the curve gives the initial (virgin) rock temperature with an accuracy believed to be of the order of 1/10°C. To make sure that drilling and blasting effects are really negligible, as any calculations which can be carried out would indicate, an occasional 14 foot hole has been run in at the bottom of a short cross-cut such as occasionally occurs in drifting. In such a hole the tests made so far indicate a steady temperature for weeks and months and go far towards removing all doubt in one's mind as to any transient effects. The agreement which has been found so far between measurements in such deep holes and the standard 7 foot ones is good, but much more work remains to be done along this line before complete confidence can be placed in the results, however satisfactory they appear to be as a whole.

Results

The results to date are shown in Fig. 2. The points determined in the present work are seen to lie very well on a smooth curve which is almost a straight line. When this is continued to the surface level, making such use as seems justifiable of the Agassiz measurements, and especially of the point[4] indicating average surface temperature, we have the geothermal curve for this particular spot—insofar as it can be considered to be established by these measurements; it is unfortunate that there have been no recent opportunities for reliable measurements in the upper mine levels. The absence of local heat sources in this region, as well as general geological considerations, would lead one to believe that this curve will probably be found to be typical of this region, but measurements in other mines will be made as opportunities offer. A comparison with all the Van Orstrand drill-hole data for this locality must also be made. It may be noted in passing that the average gradient found here of 1°C in 59.5 meters is less than half as steep as that taken by Kelvin as the average for the whole earth, viz., 1°C in 27.76 meters.

Theoretical interpretation of results

A preliminary attempt has been made at the interpretation of this curve in terms of the previous thermal history of this region, involving particularly the time since the last ice age. In order to do this it is first necessary to determine the thermal diffusivity (thermal conductivity divided by the product of specific heat and density) of the rock material. B. O. Peirce[5] made a number of careful tests of the conductivity of specimens of rock from the Calumet and Hecla mine many years ago, but unfortunately several assumptions are involved when it comes to getting diffusivities from these results. Accordingly, a few tests on specimens brought from two regions in the mine

[4] This is the point taken by Lane (reference 2, p. 703–705) as the average surface temperature at Calumet. It was arrived at in several ways, all of which give nearly the same result and is believed to be a very good average.

[5] B. O. Peirce, Proc. Am. Acad. of Arts and Sci. **38**, 652 (1903).

have been made by a method developed by the writer[6] some years ago. The specimen of rock, which must be in the form of a slab say 6 cm thick and 25 or more cm square, is mounted in a paraffined wood frame which thermally insulates the edges, leaving only the faces exposed. A fine thermocouple is located in a small hole drilled into the center of the slab, everyth ng being

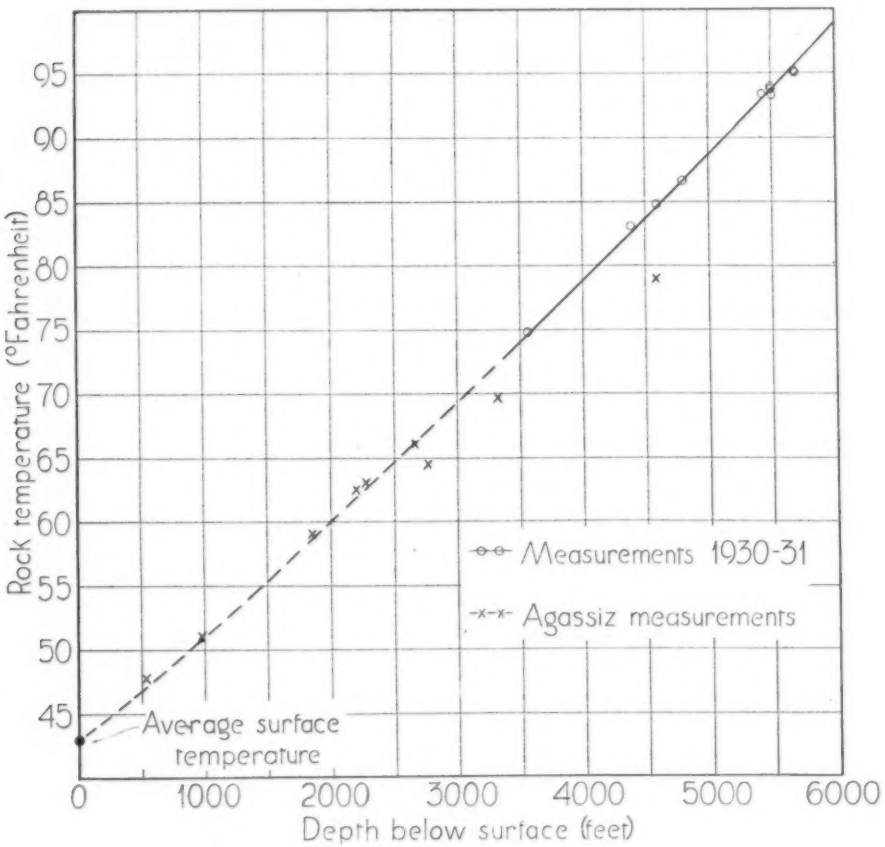

Fig. 2. Geothermal measurements in the Calumet and Hecla mines. Average temperature gradient, 0–5500 ft. depth, 1°F in 108.5 ft. or 0.00922°F/ft., or 1°C in 59.5 meters, or 0.0164° C/m. Gradient, 3500–5500 ft. depth, 1°F in 103.1 ft. or 0.00970°F/ft.

carefully waterproofed. When such a slab has been left in a constant temperature region for 24 hours and is then suddenly plunged into stirred ice water so that the faces are quickly lowered to and kept at zero, the diffusivity can be readily calculated from the time it takes for the center temperature to fall half way from its original value to the zero point. Two specimens of trap rock tested in this way gave values for the diffusivity of 0.0077 and 0.0074 respectively. Since the former specimen was from the vein and the latter

[6] For a brief account of this method see L. R. Ingersoll and O. A. Koepp, Phys. Rev. **24,** 92 (1924). For the underlying theory see Ingersoll and Zobel, "Mathematical Theory of Heat Conduction" (Ginn) p. 105 ff.

from the surrounding trap rock, which composes the great bulk of the material, the weighted average is taken as 0.0075 c.g.s. units.

With the above value, calculations[7] have been made of theoretical temperature-depth curves for some 24 different assumed "thermal histories" of this region. The simplest assumption is that the surface temperature has been about as at present for 10,000 years and that previous to this the surface was covered with ice. In other assumptions the time has been changed and in still others several glacial overflows are taken account of. The periods of time have been varied between wide limits. The equation for calculating the points of these curves is

$$\theta = \frac{x}{2h(\pi)^{1/2}} \int_0^t F(\lambda) e^{-x^2/4h^2(t-\lambda)} \cdot (t-\lambda)^{-3/2} d\lambda$$

where x is the distance below the surface, at which the temperature θ is to be calculated, h^2 the diffusivity, and $F(\lambda)$ the surface temperature function.

These curves having been calculated and corrected for the general average gradient slope, a careful comparison has been made between them and the actual curve as given in Fig. 2. The results are a little disappointing at first as the curve is so nearly a straight line anyway and so much depends on the justification for certain assumptions which cannot as yet be proved satisfactorily, e.g., that the diffusivity of the rock is unchanged for the first mile or more in depth. However, certain assumptions fit much better than others and point on the whole to a somewhat longer period—say 30,000 years—since the last glacial overflow for this region, than is usually assumed.

In addition to the acknowledgements at the beginning of this article, special thanks are due to Mr. James McNaughton, President of the Calumet and Hecla Company, for placing at the service of this investigation the facilities of the mines, and particularly to Messrs. H. E. Jefferson, H. S. Donald and R. F. Wilson for their care in making the measurements, as well as for many valuable suggestions.

[7] I am glad to acknowledge the assistance of the following of my students in making these long and arduous calculations: Miss M. C. Wolf and Messrs. R. O. Anderson, R. E. Erickson, R. G. Herb, W. L. Hole, R. W. Prucha and M. T. Rodine.

VELOCITY OF ELASTIC WAVES IN GRANITE

By L. Don Leet and W. Maurice Ewing
Harvard Seismograph Station and Lehigh University

(Received January 14, 1932)

Abstract

The velocity of elastic waves in granite was determined at Quincy and Rockport, Massachusetts, and Westerly, Rhode Island. The waves measured were generated by dynamite explosions. They were recorded by portable seismographs at distances ranging from fifty feet to four thousand six hundred feet. The observed velocities for longitudinal waves were:

Quincy.....16,260± 70 ft./sec. or 4.96±0.02 km/sec.*
Westerly....16,400±120 ft./sec. or 5.00±0.04 km/sec.
Rockport....16,670± 40 ft./sec. or 5.08±0.01 km/sec.
Average.....16,530± 90 ft./sec. or 5.04±0.03 km/sec.

A three-component seismograph, used only at Quincy, recorded transverse waves, the velocity of which was 8150±90 ft/sec., or 2.48±0.03 km/sec. From the two velocities determined at Quincy and the density of specimens taken from the shooting location, 2.65 grams/cm³, values for the bulk modulus, k, compressibility, β, rigidity, μ, Poisson's Ratio, σ and Young's Modulus E, were obtained as follows: $k=44\pm1\times10^{10}$ dynes/cm²; $\beta=2.28\pm0.05\times10^{-12}$ cm²/dynes; $\mu=16.3\pm0.4\times10^{10}$ dynes/cm²; $\sigma=0.333\pm0.005$; $E=43\pm1\times10^{10}$ dynes/cm². The form of the time-distance curves, straight lines through the origin, indicated that the waves did not penetrate deeply. Accordingly, the values obtained are for pressures of only a few atmospheres. The bearings of these results upon earlier investigations of the elastic constants of granite are discussed. Although direct comparisons between laboratory and field results are not conclusive, they indicate that the Adams and Williamson curve is incorrect for pressures below 2,000 megabars, and that there is no marked difference between dynamically and statically determined compressibilities of granite.

I. Location

DETERMINATIONS of the velocity of elastic waves in granite were made in three well-known granite quarrying areas, Quincy and Rockport, Massachusetts, and Westerly, Rhode Island. Within each area a zone of practically no topographic irregularity was selected. The continuity of the granite in each zone was proven by numerous quarries and outcrops. Maps of these zones appear in Fig. 1. The observing stations and shot locations shown there were mapped by plane table on a scale of 300 feet to the inch.

II. Method

The elastic waves were produced by the explosion of charges of dynamite which were always placed in contact with the granite. Sixty percent quarry gelatine was used. Their arrival at a point on the granite at any desired distance from the explosion was recorded by a seismograph placed at that point.

* The ± values given in this paper are probable errors.

The distance travelled by the waves and the time required constituted the observed data. Distances were scaled from the maps reproduced in Fig. 1.

The time plotted against the distance yielded a time-distance graph which was a straight line through the origin. The velocity, the reciprocal slope of this line, was computed by a least square solution.

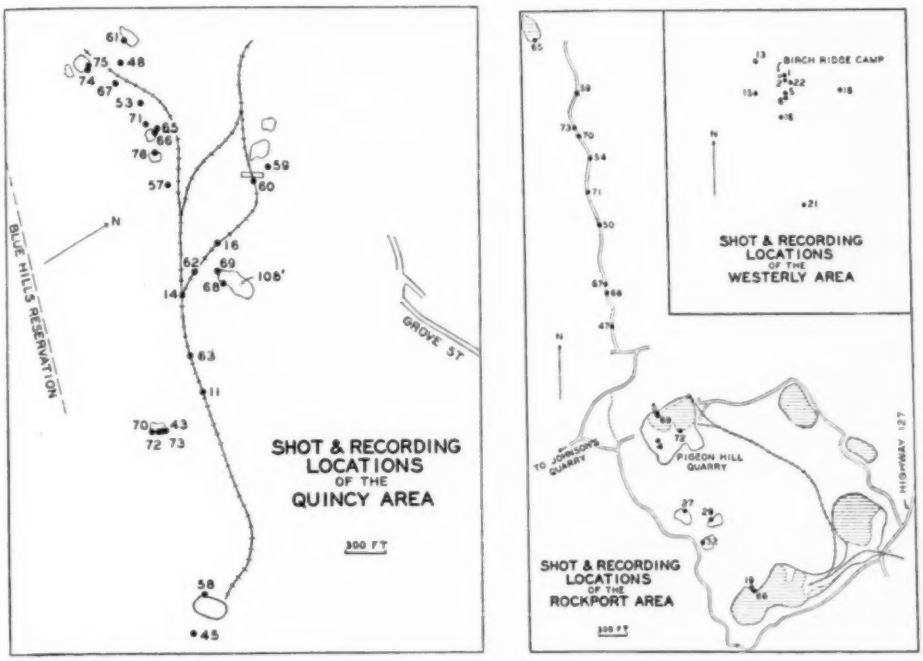

Fig. 1.

III. Instruments

The seismograph used at Quincy consisted of three components, two horizontal and one vertical, mounted on a single base. The instrument was always oriented so that one of the horizontal components recorded vibrations on a line from shot to recorder, while the other, at right angles to this, recorded vibrations transverse to that line. These are designated respectively as the longitudinal and transverse components. The longitudinal component had a free period of 0.77 sec., the transverse, of 0.62 sec., and the vertical, of 0.55 sec. Each was critically damped by a magnetic device of conventional design. The static magnification, effected by mechanical and optical means, was about one thousand.

At Westerly and Rockport, a vertical component seismograph was used. It had a free period of approximately two tenths of a second and was critically damped by a piston working in a cup of oil. The static magnification, also mechanical and optical, was about twelve thousand.

Both of the instruments recorded photographically on a strip of bromide paper one inch wide, moving at a speed of about twenty centimeters per

second. A vacuum-tube driven tuning fork controlled the interruption of a beam of light, thereby producing timing lines, which extended the full width of the record, at intervals of 0.01 sec.

The electrical circuit which was used to detonate the dynamite from the recording location also served for the registration on the record of the instant of the explosion. A modified polarized relay in this circuit was actuated both when the firing circuit was closed and when it was broken by the explosion. A mirror attached to the moving vane of the relay and in the path

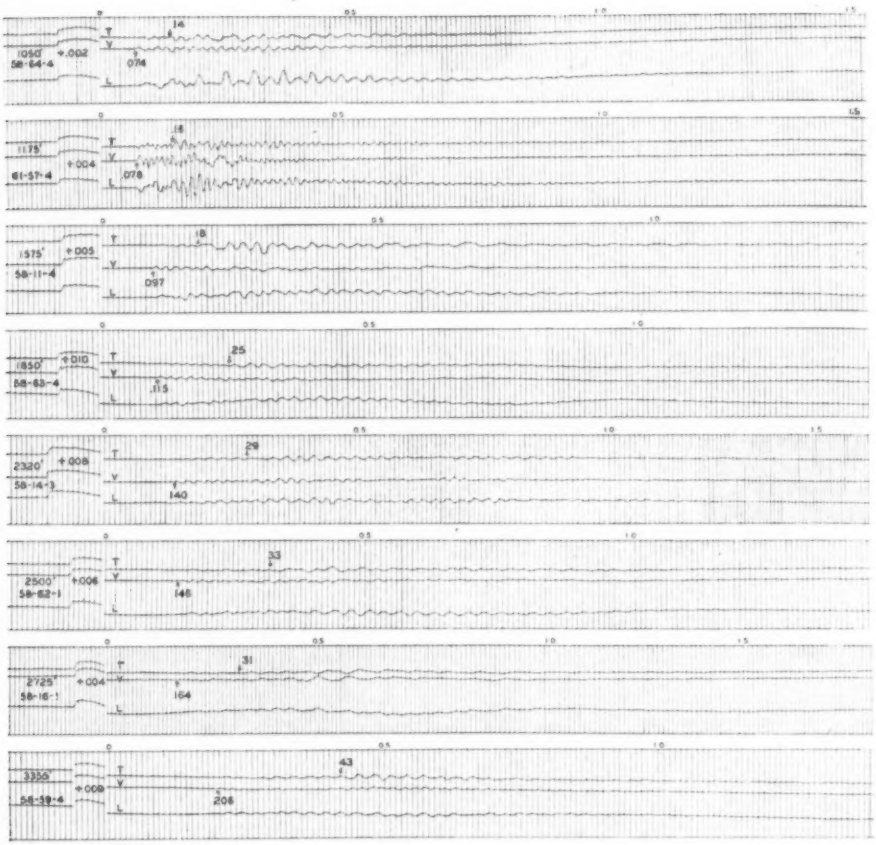

Fig. 2. Specimen Quincy records *T*, transverse component; *V*, vertical component; *L*, longitudinal component.

of the beam from light-source to seismograph effected the actual registration. In Fig. 2, the upward jog in all seismograph lines marks the deflection caused by closing the firing circuit, and the downward jog a few hundredths of a second later, the instant of the explosion.

IV. Firing Positions

At Quincy and Rockport, several abandoned water-filled quarries were used as firing locations. The charges were placed against the quarry walls and

in some cases as much as thirty feet beneath the surface of the water, and one hundred feet beneath the surface of the ground. This arrangement had the important advantages that the explosions were confined by the water and that the energy entered sound, unweathered granite. The efficient confinement of the explosions by the water made it possible to use relatively small

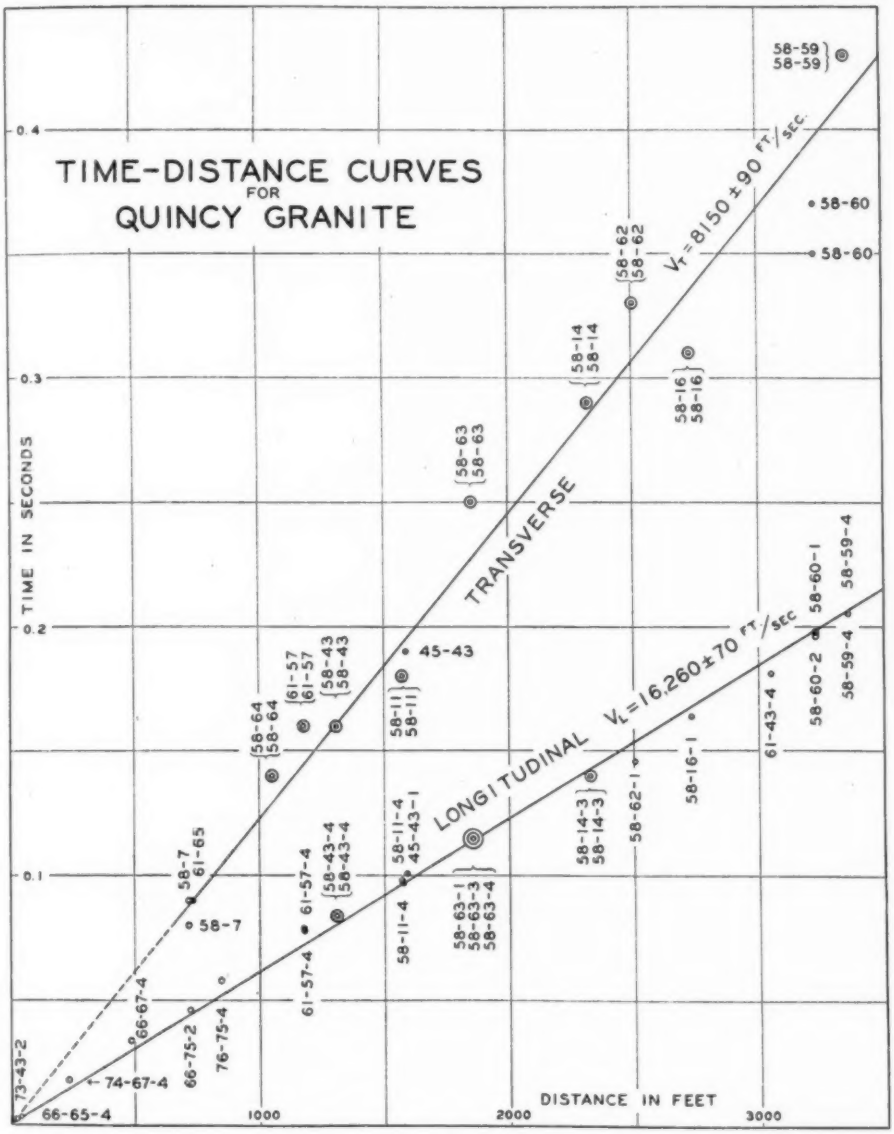

Fig. 3.

charges. About five hundred pounds of dynamite sufficed for the entire investigation.

At Westerly, shooting positions were chosen where the granite was covered with several feet of glacial drift. There the charges were placed in holes dug down to the granite.

V. Quincy Results

Specimen records from Quincy are reproduced in Fig. 2. The traces marked "T", "V", and "L" are the transverse, vertical, and longitudinal components respectively. Each record is identified by figures indicating the distance, shot position, recording position, and weight assigned it in the least square solution. For example, the first record in Fig. 2 is for a distance of 1050 feet. The figures 58–64–4 indicate that the shot was fired at position 58 (see map, Fig. 1), recorded at position 64, and accorded a weight of 4 in the least square solution. Weights were assigned on the basis of amplitude of trace and general photographic quality, which determined the accuracy with which the time could be read.

The interval between the instant of the explosion and the "0" timing line, the travel-time of the longitudinal wave on the vertical component, and the travel-time of the transverse wave on the transverse component are shown on each record.

The time-distance graph for Quincy is shown in Fig. 3. The fact that it is a straight line through the origin indicates that the waves did not penetrate deeply. It is probable that the penetration was not more than twenty meters, and certain that it did not exceed 200 meters.* The lines drawn through the points were determined by least square solutions. Their reciprocal slopes give:

$$V_L = 16{,}260 \pm 70 \text{ ft./sec. or } 4.96 \pm 0.02 \text{ km/sec.}$$
$$V_T = 8{,}150 \pm 90 \text{ ft./sec. or } 2.48 \pm 0.03 \text{ km/sec.}$$

where V_L and V_T are the velocities of longitudinal and transverse waves respectively.

The density of specimens taken from the shooting location was 2.65 ± 0.02 grams/cm.[3]

The elastic constants of the granite can be obtained by substituting these values in the following equations.[8]

$$V_L = \left(\frac{\lambda + 2\mu}{\rho}\right)^{1/2} \tag{1}$$

$$V_T = (\mu/\rho)^{1/2} \tag{2}$$

$$k = \lambda + 2/3\mu \tag{3}$$

$$\beta = 1/k \tag{4}$$

$$E = \frac{\mu(3\lambda + 2\mu)}{\lambda + \mu} \tag{5}$$

$$\sigma = \frac{\lambda}{2(\lambda + \mu)} \tag{6}$$

where k is the bulk modulus, β the cubical compressibility, μ the rigidity, E Young's modulus, σ Poisson's ratio, and ρ the density.

* This point is being investigated further.

The values obtained for these constants are:

$$\mu = 16.3 \pm 0.4 \times 10^{10} \text{ dynes/cm}^2$$
$$\lambda = 33 \pm 1 \times 10^{10} \text{ dynes/cm}^2$$
$$k = 44 \pm 1 \times 10^{10} \text{ dynes/cm}^2$$
$$\beta = 2.28 \pm 0.05 \times 10^{-12} \text{ cm}^2/\text{dynes}$$
$$\sigma = 0.333 \pm 0.005$$
$$E = 43 \pm 1 \times 10^{10} \text{ dynes/cm}^2.$$

VI. Rockport Results

Specimen records from Rockport are reproduced in Fig. 3 and the time-distance graph is shown in Fig. 5.

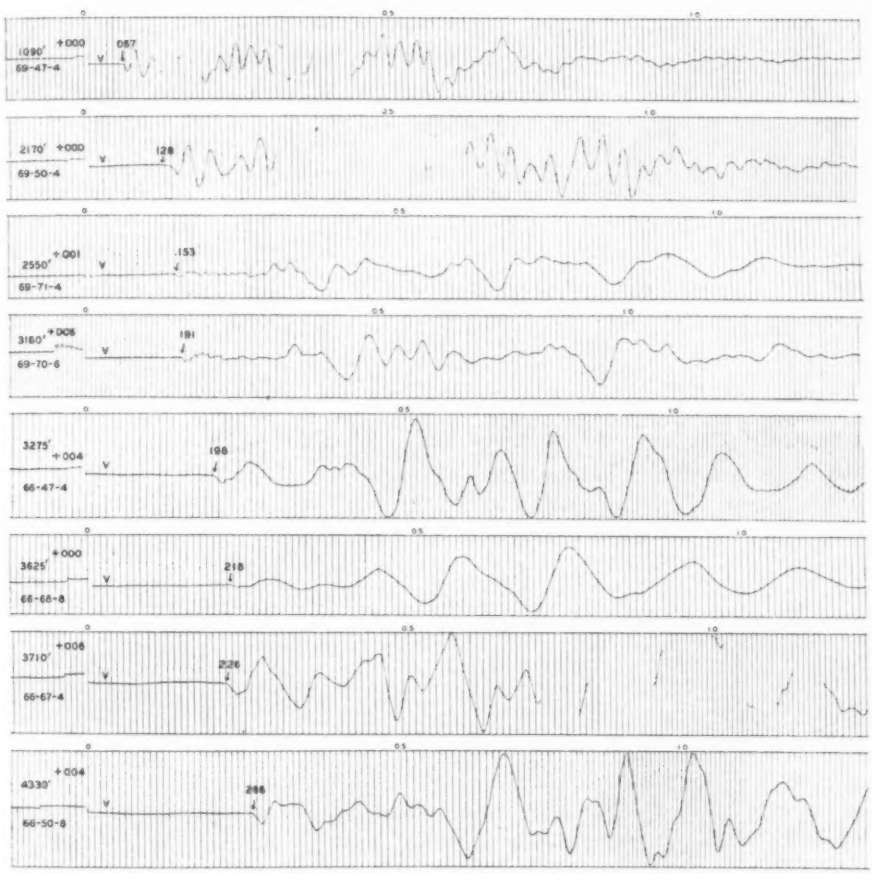

Fig. 4. Specimen Rockport records, vertical component.

The velocity for longitudinal waves obtained from these data was 16,670 $\pm$40 ft./sec. or 5.08 $\pm$ 0.01 km/sec.

VII. Westerly Results

The time-distance graph for Westerly is shown in Fig. 5. The velocity for longitudinal waves obtained from these data was 16,400 ± 120 ft./sec. or 5.00 ± 0.04 km/sec.

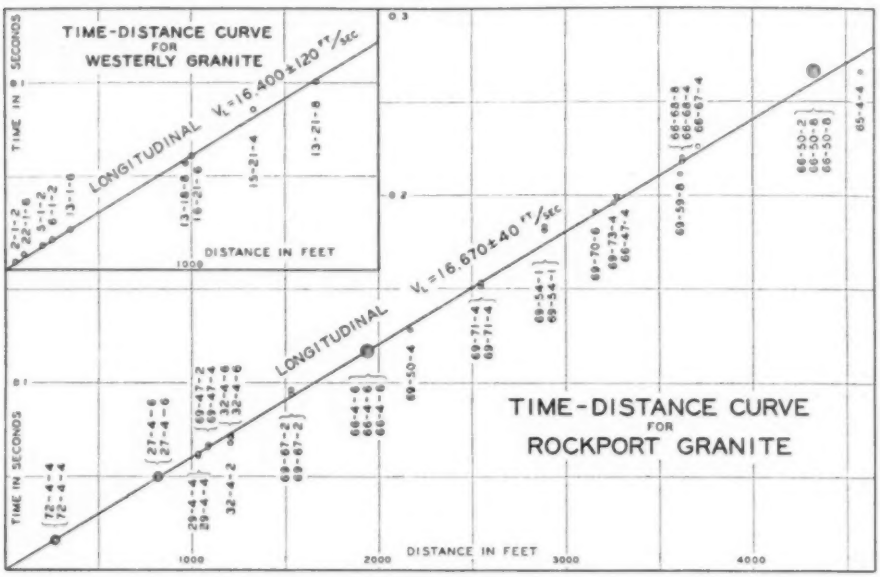

Fig. 5.

VIII. Average Longitudinal Velocity

The weighted average of these three longitudinal velocities is 16,530 ± 90 ft./sec. or 5.04 ± 0.03 km/sec.

The close agreement in velocity found here for three well-separated and different types of granite indicates that minor fluctuations in composition do not seriously affect the velocity.

IX. Analyses

Chemical

The following analyses of specimens from the regions investigated were made by Mr. Forest A. Gonyer, chemist of the Harvard Department of Mineralogy.

Mineralogical

The following mineralogical analyses were made by Mr. Chalmer J. Roy, Assistant in the Department of Geology, who was a member of the party throughout the investigation.

The Westerly analyses were made from thin sections. The Rockport and Quincy determinations were made by gravity separation methods from initial samples of 1000 grams.

TABLE I.

	Westerly		Quincy	Rockport
	Fine grain	Coarse grain		
SiO_2	74.00	70.30	72.86	74.37
TiO_2	0.10	0.31	0.18	0.18
Al_2O_3	13.86	15.10	11.74	12.88
Fe_2O_3	1.18	1.58	2.99	0.94
FeO	0.71	0.99	1.56	1.56
MnO	none	0.04	0.03	0.03
MgO	0.11	0.50	0.19	0.23
CaO	0.75	1.25	0.14	0.52
Na_2O	2.88	3.69	4.36	4.11
K_2O	6.20	5.58	5.28	4.92
H_2O-	0.09	none	none	none
H_2O+	0.31	0.40	0.41	0.31
CO_2	none	none	none	none
P_2O_5	none	none	none	none
SO_3	none	none	none	none
	100.19	99.74	99.74	100.05

TABLE II.

	Westerly		Quincy	Rockport
	Fine grain	Coarse grain		
Feldspar	45% (orthoclase)	50% (orthoclase)	60.10% (mostly Perthite)	59.48% (mostly Perthite)
Quartz	35%	35%	28.88%	35.14%
Others	20% (muscovite 15, biotite 2, magnetite 3)	15% (muscovite 3, biotite 7, magnetite 3, zircon and apatite 2)	11.02% (riebeckite aegerite fluorite titanite)	5.58% (hornblende)
	quartz and feldspar crystals range from 1/60 mm to 2 mm in diameter, average ½ mm *ca.*	quartz and feldspar crystals range from ⅓ mm to 4 mm in diameter, average 1½ mm *ca.*		

X. DISCUSSION

La Courtine

At La Courtine, Maurain, Eblé, and Labrouste[9] registered seismic waves generated by explosions and transmitted through a granitic terrane.

They recorded longitudinal waves at distances of 5.6 km, 7.7 km, and 13.9 km. When distances of this magnitude are used, a good velocity determination can be made with less accurate timing than is necessary with shorter distances such as were used in the present work. This advantage, however, is offset by the introduction of uncertainty as to the uniformity of terrane throughout the range. At La Courtine the presence of both gneiss and granite in the path of the waves lessened the certainty with which the velocities obtained may be assigned to granite.

The method of timing the blasts introduced further uncertainty. The instant of the blast, because of the failure of an electric circuit designed to register it on the records, was determined visually by an observer equipped with a chronometer. *An allowance of 0.2 sec. was made for his time of reaction.*

The mean longitudinal velocity obtained from five records in the La Courtine experiments was 5.524 km/sec. The five values, however, varied so much among themselves that they were segregated into two groups, giving averages of 5.905 km/sec. and 5.270 km/sec. respectively. The former was assigned tentatively to granite, and the latter to gneiss. An examination of their geologic map shows that this assignment is not valid because both granite and gneiss were traversed in every case.

In spite of the use of "un amortissement trés faible", the La Courtine investigators report, in addition to longitudinal waves, "autres, environ deux fois moins rapide, d'amplitude plus grande, se manifestent sur les trois composantes; nous les désignerons par le symbole '*L*' appliqué de maniére générale en sismologie aux ondes de plus grande amplitude (longues ondes). De plus, a La Courtine, station la plus rapprochée des explosions, ont été enregistrés deux fois trés faiblement sur la composante horizontale transversale, et une fois en méme temps sur la composante verticale, des mouvements trés faibles oui paraissent ainsi correspondre a des ondes a vibrations transversale de vitesse intermediaire entre celles dex deux categories précedentes; nous les désignerons par '*S*' (secondes ondes)." The use of very feeble damping must, of course, have made the identification of these later phases somewhat uncertain. In the light of experience with transverse waves from explosions, which has been accumulated since the completion of this pioneering experiment, it seems clear that the waves designated as "*L*" were actually transverse. (See, for example, refs. 3, 7). The velocity of these waves, as determined from the mean of eight observations, was 2800 m/sec., making the ratio of the velocities of longitudinal and transverse waves 1.97. This is practically the same as the value of 2.00 obtained in the present investigation.

Thus, the velocities obtained by Maurain, Eblé, and Labrouste for both transverse and longitudinal waves, 2.800 km/sec. and 5.524 km/sec., respectively, are about 12 percent higher than the values for Quincy granite. Since both the method of timing the blasts and the presence of large areas of gneiss in the La Courtine region could introduce systematic errors, this divergence is not surprising.

Comparison of laboratory and field results

There is a fundamental problem involved in a consideration of the utility of comparing small-scale laboratory determinations of rock constants with field results. For many years it has been customary to use constants determined in the laboratory for computing elastic-wave velocities characteristic of certain rocks. On the basis of such comparisons deductions have been made concerning rock types represented by velocities observed in studies of near earthquakes.[6et al] When the magnitude of the laboratory specimens in

past experiments is compared with the volume and field relationships of materials of which they are called representatives, however, this appears to be a very curious procedure indeed.

A case in point is the investigation of Adams and Coker,[1] with particular reference to the granites which they measured, since those results bear directly on the present work. Their measurements were made on granites from six widely separated localities. With the exception of Quincy, from which there were two, a single rock sample was used from each place, and from one to four test specimens were cut from each sample. The test pieces were about one inch in diameter and 3 inches long. The report states that "the rocks in all cases were air dry, having been allowed to remain in the laboratory for several weeks after they had been cut, before measurements were made". Examination of individual measurements reported shows that constants vary from specimen to specimen, and in different directions in given specimens, up to as much as 50 percent. The lateral extension curve for one Quincy specimen has a definite kink, and Poisson's Ratio at 2000 pounds of pressure is nearly 40 percent less than the average given for the specimen under pressures up to 9000 pounds. All of this raises a serious question as to the significance of the average results as representative of the Quincy granite in place.

Laboratory determinations of the constants of a sufficiently large number of samples from a given region would, of course, give a much more representative figure. It is difficult to see, however, how past determinations reported in the literature can be said to represent general averages in any sense, entirely aside from any question as to methods and accuracy of measurements. Their acceptance as a basis for deductions from studies of near earthquakes seems to imply an illusion that granite bodies, whether shallow or deep, are homogeneous masses of uniformly solid, flawless rock. Since they are anything but that, it should not seem strange if actual velocities of waves propagated through several thousand feet of granite in the field differed from those computed from the constants of isolated 1 in. $\times$ 3 in. specimens of the same granite.

Realizing this, we have nevertheless included the following comparisons on the possibility that they may record relationships between laboratory and field results which are interesting even if not particularly significant.

Relation between compressibilty of granite and pressure

Adams and Williamson[2] have determined the compressibility of several rocks and minerals by a static method at pressures of 2,000 and 10,000 megabars. In Fig. 6 a graph showing the variation of compressibility with pressure is reproduced from their report. The dotted line and marked point have been added. It can be seen that according to this graph the statically determined compressibility of granite changes very rapidly at pressures under 2000 megabars. Independent measurements of the compressibility of quartz and the feldspars which are the chief constituents of granite fail to show any such anomalous increase at pressures less than 2000 megabars, so the re-

ported behavior of granite stands out as an exception to the law demonstrated by Adams and Williamson "that at moderately high pressures the compressibility of a fresh holocrystalline rock is an additive function of the compressibility of its minerals", and resulted in their modifying it to apply only "provided the pressure is not too low". The following quotations summarize the evidence which Adams and Williamson had for this change and show the difficulty which they encountered in explaining it.

In the discussion on page 523 of their report, accompanying the graph shown in Fig. 6 here, is the following statement: "The results of Adams and Coker for six different granites provide the basis for the estimate of the range

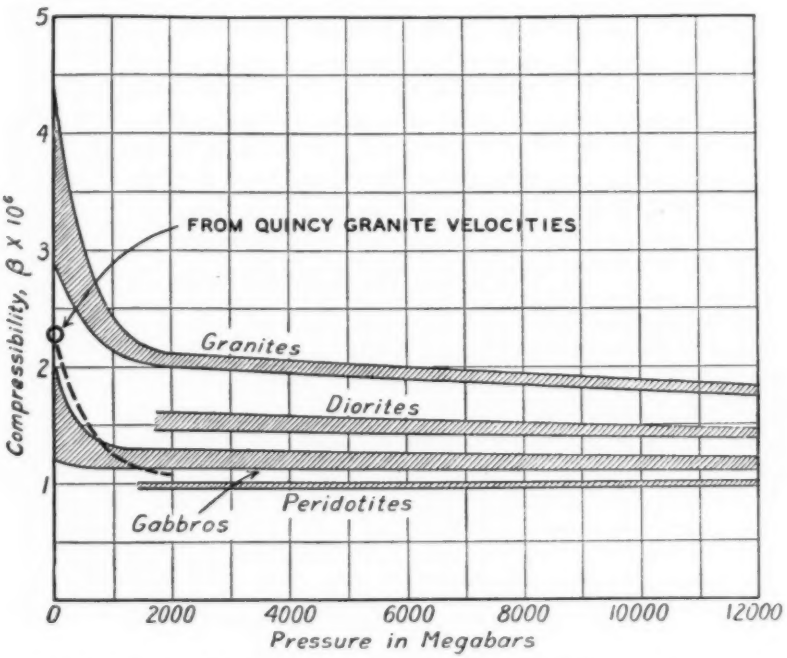

Fig. 6. Compressibility as a function of pressure, according to Adams and Williamson. The circle and dotted line have been added to their graph.

of variation for granites, at low pressures. The corresponding variation for the gabbros at low pressures is admittedly a very rough estimate representing mainly the opinion of the authors."

Thus, the only evidence for the anomalous behavior of the compressibility of rocks at low pressures seems to be the results of Adams and Coker on six granites. Even a futile attempt to find a consistent explanation for the anomaly apparently did not lead Adams and Williamson to doubt the validity of the earlier results as average figures for low pressures. Porosity was looked upon as a possible cause, but after considering it at some length, they concluded: "None of the rocks has enough porosity to explain, on the basis of detached voids or channels, more than a few percent of increase of compressibility. . . . The granites, moreover, although containing less than one

percent of pore space, show an abnormally high compressibility at very low pressures; the value at a pressure of only a few hundred megabars being nearly double that at 1000 megabars."

After suggesting two more possible explanations and citing examples both contradictory and confirmatory, they proceed to a fourth: "Very often the abnormally high compressibility at low pressures is associated with coarseness of grain—for example, the granites and marbles, which are comparatively coarse-grained, show a high initial compressibility. The Sudbury diabase, however, is moderately coarse-grained and does not show much of this effect, while the serpentine, which is very fine-grained, shows a considerable decrease in compressibility and this tendency to decrease persists even at high pressures".

In this same connection, it is interesting to note that Adams and Coker[1] had investigated the effects of grain size, including a granite in the trial runs, and concluded: "It will thus be seen that there is no correspondence between the coarseness of grain and the magnitude of the variations in the readings obtained."

This situation, which clearly proved rather difficult, and hinges essentially on the high values for the compressibility of granite obtained by Adams and Coker, is left by Adams and Williamson as follows: "About all that can be said concerning the way in which the compressibility of rocks changes with pressure is: For pressures above 2000 megabars the compressibility does not change very much; at pressures of only a few hundred megabars the compressibility is likely to be notably higher; and that the change of compressibility at low pressures, while connected with the admixture of minerals of different compressibility and with a looseness of structure existing in some coarse-grained rocks, cannot with certainty be predicted in advance."

The compressibility of Quincy granite, 2.29×10^{-12} cm^2/dynes, reported earlier in this paper, has an important bearing on the question under discussion. The issue is somewhat confused by the fact that some writers contend that high-stress, statically determined compressibilities may be considerably different from those determined dynamically, which involve low stresses. If, for the moment, we assume the equivalence of statically and dynamically determined values, it is at once apparent that the compressibility of Quincy granite as plotted in Fig. 6 agrees with the high-pressure values of Adams and Williamson and eliminates the anomalous change at low pressures advocated by them. If, on the other hand, we deny this equivalance, it seems that it is almost inescapable to assume a variation in dynamically determined compressibility similar to that shown by Adams and Williamson. The dotted curve in Fig. 6 represents this assumed variation. According to this curve, the seismically effective compressibility at 2000 megabars would be roughly one half that observed at Quincy for pressures of a few megabars. Barring compensating changes in density and rigidity, such a decrease in compressibility would give a velocity of approximately 7.5 km/sec. for longitudinal waves in granite at 2000 megabars pressure. This is far in excess of the figure 5.6 km/sec. observed by Jeffreys[6] and others as a

maximum velocity in the "granitic layer" to depths where the pressure is greater than 2000 megabars. We are accordingly forced to conclude that either the "granitic layer" is not granitic, or granite does not exhibit such rapid changes in compressibility at pressures below 2000 megabars. It appears that, whether or not we consider statically and dynamically determined compressibilities equivalent, the observed compressibility of Quincy granite is not comparable with the low-pressure part of the Adams and Williamson curve.

The good agreement between the compressibility measured at Quincy by a dynamic method and that measured statically by Adams and Williamson suggests that the marked discrepancy between static and dynamic values which has been reported[4,5] does not exist for granite. It should be noted that though the only quantities which they measured were the compressibility and density, on page 520 of their paper they have computed the velocities of seismic waves in the rocks they studied. The values which they show for the velocities depend upon the value 0.27 which was assumed for Poisson's ratio. If the value 0.333, which was determined at Quincy, were substituted in their computations, the resulting velocities would be in agreement with those determined in the present investigation.

XI. Acknowledgments

The investigation was conducted by the Harvard Seismograph Station during the summer of 1931. The instruments used were built and patented by H. Gordon Taylor in the name of the Geophysical Exploration Company. They were purchased through generous gifts from Mr. Edward Mallinckrodt, Jr., chairman of the Committee to Visit the Division of Geology at Harvard, and Professor R. A. Daly. The cost of the balance of the field equipment and all operating expenses were met by a grant from the Milton Fund of Harvard University.

We are indebted to the Atlas Powder Company for permission to shoot on their property at Quincy, to the Rockport Granite Company for permission to use their property at Rockport, and to Mr. John H. Austin for permission to run our Westerly profile on his property. Their ready cooperation was an important factor in making possible the rapid and successful completion of the program.

Bibliography

[1] F. D. Adams and E. G. Coker, "An Investigation into the Elastic Constants of Rocks, More Especially with Reference to Cubic Compressibility". *Carnegie Institution of Washington Publication No. 46*, 69 pages (1906). Washington, D. C.

[2] L. H. Adams and E. D. Williamson, "The Compressibility of Minerals and Rocks at High Pressures", *Journal of the Franklin Institute*, **195**, 475–529 (1923). Reprinted as No. 484 of *Papers from the Geophysical Laboratory*, Carnegie Institution of Washington, Press of J. B. Lippincott Company, (1923).

[3] B. Brockamp and Mothes, H., "Seismische Untersuchungen auf dem Pasterzegletscher", *Zeitschrift für Geophysik* **6**, 485 (1930), Göttingen.

[4] R. A. Daly, "The Effective Moduli of Elasticity in the Outer Earth-Shells", *Gerlands Beiträge zur Geophysik* **19**, 194–209 (1928).

[5] R. A. Daly, "Nature of Certain Discontinuities in the Earth", *Bulletin of the Seismological Society of America* **20,** 41–52 (1930), Stanford.

[6] Harold Jeffreys, "The Earth: Its Origin, History, and Physical Constitution", 2nd Edition, Cambridge University Press, 346 pages, 3 plates, 16 figs., index (1928). Price $5.50, Cambridge.

[7] Walter Korte, "Beiträge zur experimentellen Seismik", Zeitschrift für Geophysik **7,** 64, (1931), Göttingen.

[8] A. E. H. Love, "A Treatise on the Mathematical Theory of Elasticity", Published by Cambridge University Press, Fourth Edition, 643+xviii pages (1927). Price 40 s. net.

[9] Ch. Maurain, L. Eblé, and H. Labrouste, "Sur Sismiques des Explosions de la Courtine", *Journal de Physique et la Radium* **6,** 65–78 (1925).

ASYMMETRY OF SOUND VELOCITY IN STRATIFIED FORMATIONS

By Burton McCollum and F. A. Snell
McCollum Exploration Company, Houston, Texas

(Received February 23, 1932)

Abstract

In the course of explorations of subsurface geology by the seismograph the authors have frequently observed the pronounced effect of stratification on the velocity of seismic waves in shales, and this effect has often been utilized in practical seismography. Recently an opportunity was afforded for securing additional quantitative data on the velocity normal to and parallel to the bedding planes. The paper points out that the velocity parallel to the planes of stratification is, in some instances, as much as fifty percent higher than the velocity in a direction normal to the bedding planes. It is shown also that inclined stratified beds exhibit a higher apparent point-to-point velocity when sound travels in an up-dip direction than when traveling down-dip. The paper describes a procedure whereby this effect may be utilized for determining the direction and approximate magnitude of the dip in such stratified deposits. The method has proved to be of considerable practical importance where the stratified formations are obscured by overlying deposits.

General Outline

ONE of the earliest applications of the seismograph was its use in profiling salt domes, limestone horizons, or other high velocity rock formations buried beneath sands, clays, or shales of lower velocity. In this method of profiling, the portion of the sound wave utilized travels substantially along the path shown diagrammatically in Fig. 1. As is well understood, the time interval between the arrival of the wave at two stations some distance apart

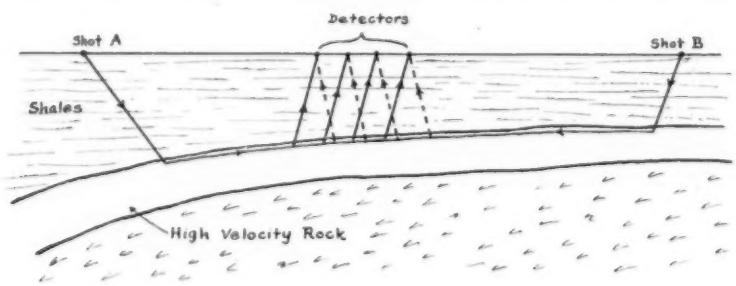

Fig. 1. Characteristic geology for diffraction profiling.

along the line of propagation of the wave will be dependent on whether the wave is being propagated upslope or downslope. By measuring this velocity asymmetry the direction and magnitude of the slope of the high velocity bed can be determined.

Commonly spoken of as diffraction or refraction shooting, this method

of profiling found wide application in many parts of the world, and still possesses much merit as a means of subsurface investigation in many localities. Its technique is now the most elementary part of the stock-in-trade of the oil seismologist. The procedure for obtaining a complete cross-sectional profile usually consists in shooting a line of detector stations from both directions. Generally the overall time of travel of the sound wave is recorded from each shot to its appropriate stations, and complementary time-distance curves are plotted for the series of shots from each direction.

Whatever may be the subsequent steps in making use of the primary time-distance data, all the calculations must necessarily be based on the deviation,

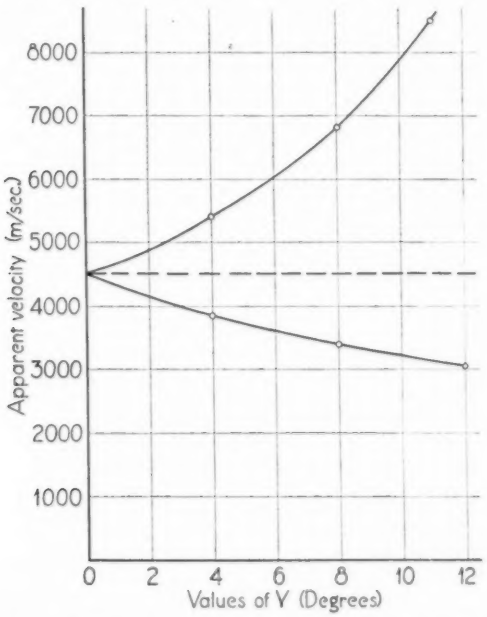

Fig. 2. Chart for showing relationships between apparent velocities when shooting upslope and downslope. Y = angle of slope; V_1 taken as 2500 m/sec.; V_2 taken as 4500 m/sec.

or asymmetry, of the apparent velocity relative to the true velocity of the rock which is being profiled. This asymmetry of the upslope and downslope values is best shown by comparing the station-to-station velocities derived by dividing the horizontal distance from one station to the next by the interval time, Δt. Fig. 2 shows the striking asymmetry which results from the slope effect in an ordinary case, while in salt dome work it is well-known that infinite apparent velocities are not uncommon.

The above described method, as is well understood, gives the slope of the *contact surface* between the high velocity bed and the overlying low velocity deposits. In many cases this contact surface, where it has not been subject to irregular erosion, will conform to the structural outlines of the subsurface geology. In such cases the simple method above outlined is adequate to give the desired information on the subsurface geology. There are many cases, however, where this situation does not exist, and in such cases the method

fails to reveal the subsurface picture. Fig. 3 shows a typical case of this geology. Here the shale, 1, in which there exist more or less definite bedding planes, has been folded into a definite structural relief. Subsequent to this folding, extensive erosion took place leaving a surface shown by the line 2–3, and on this surface later deposits, 4, were laid down. In general, under these conditions, the shale, 1, practically always has a higher velocity than the recent deposits, 4. However, if we attempt to apply the simple method above described, we shall derive not the structural picture but merely a poor approximation to the erosional surface between the shale and the recent deposits. To further increase the difficulty there is often no definite reflecting

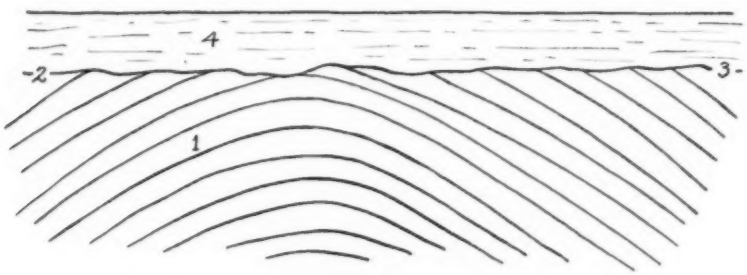

Fig. 3. Anticline hidden under unconformable beds.

surface at greater depth which can be used to define the structure by reflection shooting. Under such conditions the true structural picture cannot be revealed by any method of subsurface exploration heretofore published. These conditions are very frequently encountered and the purpose of this paper is to describe a method of exploration that is often capable of giving a correct structural picture where this situation exists. This method has been called "dip shooting" to distinguish it from the slope shooting heretofore used by geophysicists.

Origin of Dip Shooting

The method of profiling stratified formations which this paper outlines does not supercede any other method, but is simply a special development which is useful and applicable in certain cases which cannot be handled in other ways. The method is based upon an asymmetrical velocity effect which is analogous to the slope asymmetry which forms the foundation of refraction profiling. This "dip effect", as it may be called, has been observed and utilized in places as widely separated, geographically and geologically, as Canada and Venezuela. So far as our observations show, the phenomenon is common to most stratified rocks and even to some which are not ordinarily classed as stratified.

Although the dip effect is analogous to the asymmetry produced by slope, it may arise from a different cause. The velocity asymmetry associated with stratification is often due to the effect of the bedding planes, dipping with respect to the surface, and not to a sloping subsurface.

On this fact depend two important corollaries: First, the method neces-

sarily gives a structural rather than a topographic profile; for whereas the refraction method yields a structural profile under certain conditions, it is obvious that in the case of an eroded and buried land surface, for example, it merely records the old topography. Second, the method does not presuppose the existence of any cover over the formation being investigated. In this case investigation with the seismograph sounds rather paradoxical, but the condition, nevertheless, is a real one and is frequently met with in tropical work. For example, shale areas which have been folded and subsequently leveled may be covered with just sufficient surface soil to make pitting difficult and expensive while heavy vegetation obscures all the outcrops. Here the seismograph, using the dip method, can profile the structure or work out the contours with almost as much facility as surface geology displays in exposed regions. On the other hand, the presence of considerable recent overburden does not interfere with the use of the method.

A third point which may be noted here as bearing on the theory of interpretation of the records, which we shall refer to later, is that the *refraction* system of profiling does not concern itself with the sound wave as it travels through the rock being profiled, but only with its emergence into the overlying formation. On the other hand, the dip method must take into account the phenomena attending the travel of the wave at a considerable depth within the body of the stratified rock itself.

Evidence of Velocity Asymmetry

The idea that stratification might influence velocity in a way to be of practical value was first suggested to the authors by J. E. Brantly. It required considerable field work and the accumulation of extensive data, however, before it was possible to sift the velocity variations due to stratification from the variations due to several other causes. Some of these other causes will at once occur to anyone who has used the seismograph. Among them we might list such factors as varying velocity in the overburden, dip in the overburden compounding with the dip effect from the subsurface, differences of surface elevation, slope of the subsurface as distinguished from its dip in cases where the subsurface was eroded before being buried, etc.

After making due allowance for such contributory causes as the foregoing, however, it was found that the sound wave travels faster parallel to the bedding planes than it does perpendicular to them.

There are two lines of evidence to show that this is the case. The first is direct velocity measurement both parallel and normal to the bedding. Measurements of this kind were carried out recently on some exposures of the Lorraine Shales in the Province of Quebec, Canada. These shales have been highly folded along the Lake Champlain-Quebec Fault, and in a number of places are to be found standing on edge. This afforded an opportunity to measure the velocity perpendicular to the laminations of the shale and compare it with the velocity of the same formation in a horizontal position nearby. The following are some of the velocity readings.

TABLE I. *Showing velocity of sound parallel and normal to bedding.*

Velocity parallel to bedding planes. (Meters per sec.)	Velocity normal to bedding planes. (Meters per sec.)	Ratio
4450	3205	1.39
4655	3310	1.41

From the above figures it will be seen that the velocity parallel to the bedding is in the neighborhood of 40 percent higher than it is in a direction at right angles.

The second type of evidence which shows the existence of this asymmetrical effect is the difference in apparent interval velocity found in shooting up and down the dip of stratified rocks which have been tilted out of the horizontal position. Before presenting the field observations which illustrate this aspect of the problem, let us consider what theoretical explanation can be advanced to account for velocity variations as produced by dip.

It is admittedly very difficult to make an exact analysis of wave paths in a non-homogeneous medium like shale, and we do not attempt to offer any comprehensive solution here. Even a preliminary study of the problem, however, suggests that in addition to the refraction produced by more or less alternating layers of different velocity, there is also refraction due to increase of velocity with depth in most cases. Furthermore, it seems logical to suppose that considerable absorption of the wave energy would occur when the wave path cuts across the bedding, and that this absorption would tend to disappear when the path parallels the stratification.

Consideration of these and similar factors affecting the travel of the sound wave leads to the conclusion that we can make a useful approach to the problem by studying some of the simpler typical subsurface conditions such as are often met with in the field. From these simple cases it should be possible to extend the analysis to more complicated conditions with the aid of further experimental work. It is to be hoped, therefore, that additional data bearing on the subject will be forthcoming from localities where the dips are more accurately known than in many areas which the authors have investigated.

To return, however, to those instances where it can be readily seen that dipping strata would give rise to asymmetrical velocities, the three clearest cases, perhaps, are the following:

CASE 1

The first example in which dip will obviously give rise to velocity asymmetry is illustrated diagrammatically in Fig. 4. Almost any stratified formation has occasional hard layers which give a somewhat higher velocity than the average. These layers may be of the nature of "markers" which are readily distinguished by some peculiarity, or they may be almost unrecognizable to the eye while still possessing abnormal velocity characteristics. Such layers act as conductors for the sound wave because they offer the path of least time.

In the diagram, Fig. 4, layers of the above-the-average velocity are represented at *a-a*. In shooting a line of stations in an up-dip direction, the time intervals between stations 1-2 and 2-3 will be short, while between stations 3-4 and 4-5 the intervals will be longer, because these stations are beyond the zone of influence of the fast layer "*a*". However, on advancing farther down the line, other laminations of higher-than-average velocity will be encountered and the intervals will again be shortened. Conversely, when shooting down-dip, many of the intervals will be lengthened by these same layers.

Now it is expectable that the flatter the dip (until it becomes nearly horizontal), the more continuous will be the effect on the time intervals. Therefore, moderate dips, such as are usually found in practice, will show this type of velocity asymmetry to the best advantage, and those formations which exhibit the least uniform stratification will generally be the ones whose dip it will be easiest to determine.

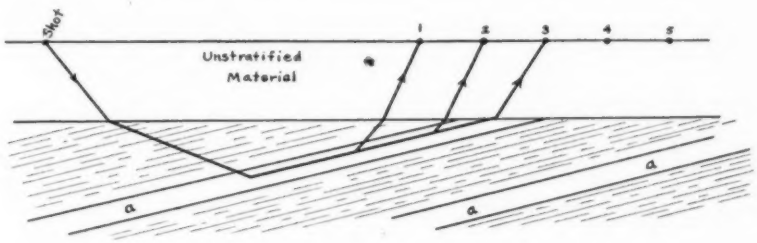

Fig. 4.

It may be objected that the foregoing analysis is based on the slope-effect and not on dip. To a certain extent this is true. In extreme cases, like that shown in Fig. 4 where the high velocity layers are unusually thick and clearly defined, the interpretation can be based purely on the theory of refraction shooting. But in a majority of cases the layers of higher velocity are relatively thin, rather closely spaced, and seldom recognizable in the outcrops. Thus it seems to be the aggregate effect of many sloping layers which generally produces asymmetrical velocities such as those tabulated in the latter part of this paper. Consequently, there is no choice but to make the interpretation in terms of dip rather than in terms of slope.

Case 2

In Fig. 5a the wave is shown traveling in an up-dip direction. Exaggerating conditions somewhat for the sake of clearness, we may assume that the wave, taking the path of least time, descends fairly deeply into the shale and is steadily refracted on a curved path until the critical angle is reached, when it follows the bedding and finally emerges at the surface giving the time interval Δt_1. In shooting down dip, the principle of reversibility of wave paths tells us that, between identical surface points, the path will be the same from either direction and there can be no difference in the recorded times. Similarly, if the subsurface dip is everywhere uniform, equal shooting distances will give identical paths even though the surface points are not the same.

From this it follows that a perfectly uniform dip, in this sense, will not result in any velocity asymmetry.

On the other hand, if the dip is non-uniform as illustrated in Fig. 5 and as is most commonly the case under actual conditions, especially close to an anticline, the wave paths will not be the same provided the surface points are shifted. Thus, in Fig. 5b, the shooting distance, L, is the same as in 5a

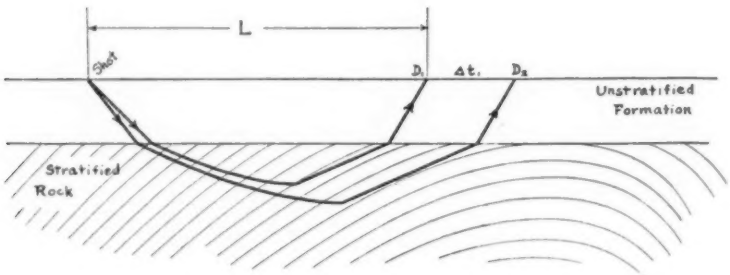

Fig. 5a.

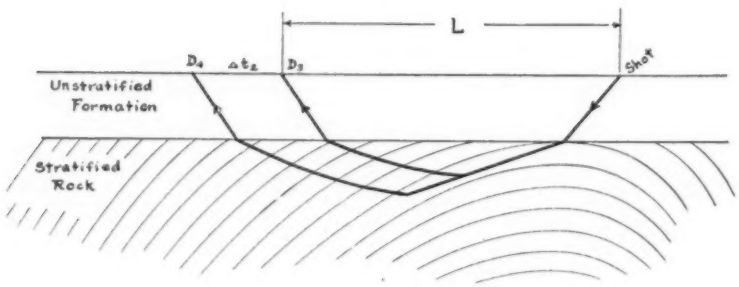

Fig. 5b.

but the setup is shifted slightly so as to intercept different conditions of dip. Consequently, the wave paths are not the same for the two cases, and both the overall times and the station-to-station time intervals will be different. In other words, Δt_2 will not, in general, have the same value as at Δt_1, and the apparent velocities will show an asymmetrical effect.

Case 3

A third type of velocity asymmetry is often due to a relationship which frequently exists between the hardness and the age of rocks and shales. In shooting up-dip (except in the case of overturned folds), the wave is traveling from younger to older beds. Now in many cases it has been found that the hardness of rocks increases with their age and the velocity closely corresponds to the hardness. Consequently, in shooting from a given shot point to a detector setup located in an up-dip direction, the tendency is for the velocity to increase, and the converse is true in shooting down-dip. This tendency, while by no means invariable, must nevertheless be taken into consideration as a causative factor. Whenever this effect exists it gives a true indication of the direction of the dip of the bedding planes.

The reader's attention has already been called to the complex nature of the problem under discussion. It is well to emphasize again, however, that many factors affecting the velocity of sound in stratified rocks have not yet been mentioned. Indeed, many factors doubtless exist which have not yet been recognized. Therefore, there is still room for much experimental investigation along these lines. At the present time, lacking a theoretical explanation which covers all phases of the phenomenon, the existence of the dip-effect and the validity of its use as a working method are best evidenced by numerous field observations; a few of which follow.

Examples of the Dip Effect

Table II. *Material: Tertiary shale in Venezuela. Shooting distance: 1500 meters. Station spacing: 250 meters.*

NW Shots		SE Shots	
Stations	Velocities	Stations	Velocities
112–111	2450	106–107	2660
111–110	2360	107–108	2720
110–109	1980	108–109	3120
109–108	2080	109–110	2600
108–107	2320	110–111	2500
107–106	2400	111–112	2720
106–105	2400	112–113	2200
105–104	2600	113–114	2180
Average	2324		2588

The velocity values shown in Table II were recorded in the course of surveying a line which was laid out approximately at right angles to the general strike of the area. The total distance covered by the stations is two kilometers, and the stations have been so chosen as to represent the same portion of the subsurface in the northwest and southeast shot groups, after making allowance for the probable amount of throwback of the reference point. The shale, in this case, was covered by clay, but the evidence of various nearby test wells pointed to the cover being fairly uniform in thickness. Consequently, the difference in the velocity averages indicates a regional dip to the southeast. The variations of the individual velocity values show the effect of local folding.

We have already mentioned the fact that some relatively unstratified formations show asymmetry due to dip. The following data were gathered in areas of indistinctly bedded sands and clays in Western Venezuela.

The values in Table III cover a distance of ten kilometers with an allowance for throwback as in the case of Table II. The difference between the average velocities is small, but in view of the circumstances a larger difference would hardly be expected. These shots were taken at a distance of about twenty kilometers out from the foothills of the Perija mountains. The dip represented by the averages is simply the normal depositional dip of water borne sediments laid down along the mountain front. The dip is toward the

TABLE III. *Material: Tertiary sandstones, sands and clays. Shooting distance: 6000 meters. Station spacing: 500 meters.*

E Shots		W Shots	
Stations	Velocities	Stations	Velocities
165–163	3290	127–129	3420
163–161	2980	129–131	2980
161–159	3120	131–133	3090
159–157	3120	133–135	3050
157–155	3050	135–137	2970
155–153	3420	137–139	2990
153–151	3380	139–141	2940
151–149	3160	141–143	2910
149–147	3050	143–145	3010
147–145	3900	145–147	2600
145–143	2940	147–149	3730
143–141	3210	149–151	2870
141–139	3290	151–153	3470
139–137	3180	153–155	2910
137–135	3080	155–157	3120
135–133	3090	157–159	2980
133–131	2930	159–161	3290
131–129	3080	161–163	3160
129–127	3080	163–165	3200
127–125	3290	165–167	3330
Average	3182		3101

east, that is, away from the mountains, and would doubtless be steeper if investigated nearer the foothills. To illustrate this latter point we may average the western half of the group of velocities, Stations 125 to 147, with these results: East shot average, 3117 meters per second; West shot average, 2996 meters per second; showing a greater difference than for the whole group.

TABLE IV. *Material: "Mottled clay", Venezuela. Shooting distance: 500 meters. Station spacing: 500 meters.*

E Shots Apparent velocities	W Shots Apparent velocities
2840	2480
2870	2480
2340	2500
2530	2250
2810	2290
2400	2420
2780	2170
2600	2140
Average 2646	2341

Table IV shows the results of some scattered shots on two lines at right angles to the Perija Mountain front. The material is what is known as the "Mottled clay" of the Lake Maracaibo region. It is a compact clay with very

little admixture of sand and has no recognizable bedding. In the locality where the above data originated, however, there were good reasons to suppose that the regional dip was eastward, in a direction away from the hills. Drilling subsequently indicated that such was the case for the underlying formations, but the seismograph results appeared to be the only definite evidence that the same dip existed in the mottled clay cover. Shots at 1000 meters distance confirmed the figures obtained from the 500 meter shots, as the following tabulation will show.

TABLE V. *Material: "Mottled clay", Venezuela. Shooting distance: 1000 meters. Station spacing: 500 meters.*

	E Shots Apparent velocities	W Shots Apparent velocities
	2320	2360
	2660	2550
	2780	2550
	2690	2580
	2940	2580
	2800	2580
		2230
		2170
		2320
Average	2698	2436

TESTS FOR KNOWN STRUCTURES

Perhaps the best confirmation of any geophysical theory is its test on known structures. We have carried out several tests of this sort with very gratifying results. One of the earliest, and at the same time, one of the most comprehensive trials of the dip method of shooting was made on a group of

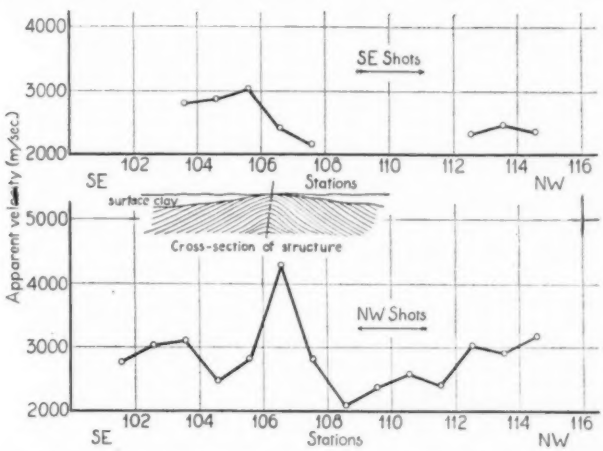

Fig. 6. Dip-velocity curves across structure miocene shale—Venezuela.

shale structures, arranged in echelon, near Maracaibo, Venezuela. These folds, outcropping through the mottled clay cover, had been worked out by surface geology, and their axes fairly well located. The seismograph was

employed to check the location of the recognized folds and to search for others which were not exposed on the surface.

The results tallied with the geological picture in a manner which could hardly be questioned. It is not possible, here, to present all the data which were gathered, but one example may not be out of place to illustrate this special application of the method to exposed or very thinly covered folds which have been planed by erosion before being buried. Fig. 6 shows the apparent velocities given by shots from opposite directions plotted in the form of velocity curves. The line of survey crossed one of the known shale structures already mentioned. The cross-section of the fold is shown diagrammatically at its proper location near Station 106. The northwest flank of the anticline is considerably steeper than the southeast flank. In shooting from the NW, note how the velocity rises to an abnormally high value due to the steep dip of that flank. A corresponding minimum value is given by the southeast shots, and the reverse occurs in both cases on the other side of the fold after crossing the crest. Doubtless a part of the velocity asymmetry is due to slope effect on the surface of the shale under the clay cover, but repeated observations seem to demonstrate that this will not account for all of the velocity variation which is found in such cases.

Application of Dip Shooting

The main essential in utilizing the dip effect is to shoot all stations from opposite directions. Generally the stations are set along a straight line which is laid out, as nearly as may be, at right angles to the strike of the area in order to get the maximum effect from the dip. Shooting distance, station spacing, etc., must be determined according to what is known of the surface and subsurface conditions, as with any other method. Often special preliminary shooting is necessary to reveal this information.

It is possible that satisfactory results might be obtained from the use of overall velocities, derived from the total time of travel of the wave from shot to detector, but the authors have always preferred to employ station-to-station velocities because they not only localize the dips better, but also bring out the velocity asymmetry more sharply.

In accordance with this practice, then, the interval velocities given by the shots from both directions may be tabulated and the complementary values compared. It is necessary to make allowance for the throw-back of the reference point as is customarily done in all profile work, although here the exact amount to allow for is indeterminate on account of the variable form of the wave path. The fact that the wave may penetrate to an appreciable depth in the shale, and that all parts of the path are generally curved, renders it impossible to do more than make an approximate calculation of the throw-back, but this will usually suffice for practical purposes.

The velocity asymmetry observed on any setup may, of course, be due to two causes. First, the slope of the contact between the shales and the overlying formations, and second, the velocity asymmetry in the shales themselves. In order to determine and eliminate the former, it is usually best to shoot a

correction shot on each setup from a point symmetrically placed with respect to two detector stations, as, for example, at a point midway between them. If the spacing of the stations has been properly chosen with respect to the local geological conditions, which can always be accomplished, the required correction can easily be made.

The most difficult phase of the work is to translate the velocity readings into terms of degrees of dip. There seems to be no theoretical basis for a rigorous mathematical treatment which would accomplish this. Consequently, the best solution is probably an empirical calibration of the velocity values on some structure where the dips are accurately known. In most cases, however, it is not of great importance to know the absolute dip. For general profiling and reconnaissance the direction and relative steepness of the dips is all that is required. The use of relative values permits comparison of the size of different structures, shows which side of the fold is steeper, and, of course, locates the crest of the fold as accurately as if precise quantitative dips were obtained.

Summary

A large amount of accumulated data similar to that herein presented demonstrates that the method described above provides a satisfactory means of outlining structure in stratified and other asymmetrical deposits where conditions are such that the usual methods of refraction and reflection cannot be applied.

THE CALCULATION OF THE MOTION OF THE GROUND FROM SEISMOGRAMS

By H. A. Wilson
Rice Institute, Houston, Texas

(Received January 30, 1932)

Abstract

The equation of motion of a mechanical seismograph is $-\ddot{x}=\ddot{y}+2k\dot{y}+p^2y$ where x is the ground displacement and y the seismograph deflection. This equation may be solved for y when x is supposed known or for x when y has been observed as a function of the time. In this paper both of these ways of solving the equation are considered. The motion of the seismograph due to a train of waves starting at $t=0$ is considered and also the motion due to the arrival of a single wave. In each case seismographs with several periodic times and either undamped or critically damped are considered. Curves are given showing the motion of the ground and the calculated motion of the seismograph. The motion of the ground corresponding to several simple assumed seismograms is also worked out and shown by means of curves. The motion corresponding to a given seismogram depends greatly on the periodic time and damping of the seismograph. Finally the ground motion is deduced from two actual seismograms due to dynamite explosions. An integraph is described which enables the calculations to be done more quickly.

IN THE case of seismograms obtained with mechanical seismographs having a definite periodic time and damping coefficient the motion of the ground can be deduced from the seismograms without serious difficulty. This of course is well-known, but so far as the writer is aware few if any results of such calculations have been published, at any rate for seismograms obtained in geophysical prospecting.[1]

The seismograph consists essentially of a heavy mass supported on springs, or in some other way, so that it is free to oscillate. The motion of the ground relative to this mass is magnified and recorded.

Consider a mass m supported by a spiral spring from a stand resting on the ground. Let y denote the upward displacement of the mass m, relative to the stand, from its equilibrium position, so that y is equal to the decrease in the length of the spring due to the upward displacement of the mass relative to the stand. When y is zero the force exerted by the spring is equal and opposite to the weight of the mass.

Let x denote the upward displacement of the ground and stand due to a wave in the ground relative to axes supposed fixed. The equation of motion of the mass m is then

$$m(\ddot{x}+\ddot{y}) = -mp^2y - 2mk\dot{y}.$$

[1] H. Arnold, Zeits. f. physik. Erdkunde **10**, 269–317, gives calculations of the ground motion from several seismograms.

Here p^2 is the restoring force per unit mass exerted by the spring when its length is increased by unity and $2k$ is the damping force or viscous resistance, per unit mass, to the motion of the mass m when moving relative to the stand with unit velocity. Hence we have

$$\ddot{y} + 2k\dot{y} + p^2y = -\ddot{x}.$$

We see from this equation that if initially $x=0$, $\dot{x}=0$ and the mass is at rest in its equilibrium position, then when the ground begins to move we shall have $y=-x$ because at the start y and $\dot{y}$ are zero so that $\ddot{y}=-\ddot{x}$. This merely means that at the start the mass m remains at rest relative to the fixed axes. But y will only be equal to $-x$ during an interval small compared with the natural period of vibration of the mass m.

The above equation may be regarded as giving y when x is known or as giving x when y is known. In previous discussions of the theory of seismographs, which the writer has seen, the motion of the ground or x has been assumed known and the resulting values of y have been calculated.

The solution of the equation $\ddot{y}+2k\dot{y}+p^2y=-\ddot{x}=F(t)$ where $F(t)$ denotes a function of the time t and when y and $\dot{y}$ are zero at $t=0$ is

$$y = \frac{1}{\mu}\int_0^t e^{-k(t-\eta)} \sin \mu(t-\eta)F(\eta)d\eta$$

where

$$\mu = (p^2 - k^2)^{1/2}.$$

By means of this equation it is easy to calculate the seismograph deflections y as a function of t for any simple values of $F(t)$.

In homogeneous ground an explosion of dynamite would be expected to produce single waves, or short trains of waves, so that it is important to consider the motion of the seismograph due to single waves. To represent a single wave we may suppose that $x=A(1-\cos\omega t)$ from $t=0$ to $t=2\pi/\omega$ and that $x=0$ when $t<0$ and when $t>2\pi/\omega$. This gives the wave shown in Fig. 1.

Fig. 1

Fig. 2.

Before considering this case it will be convenient to consider the case when $x=A(1-\cos\omega t)$ when $t>0$ and $x=0$ when $t<0$ which gives the series of waves shown in Fig. 2.

The equation $x=A(1-\cos\omega t)$ gives $\ddot{x}=A\omega^2\cos\omega t$ so that the expression for y becomes

$$y = -\frac{A\omega^2}{\mu}\int_0^t e^{-k(t-\eta)} \sin \mu(t-\eta)\cos\omega\eta d\eta.$$

We shall consider two cases. One when the seismograph is undamped so that $k=0$ and one when the seismograph is critically damped so that $\mu=0$.

In the first case with $k=0$ we get

$$y = \frac{A\omega^2}{\omega^2 - \mu^2}(\cos \omega t - \cos \mu t)$$

and in the second case with $\mu=0$ we get

$$y = \frac{A\omega^2}{\omega^2 + k^2} \cos (\omega t + \alpha) + \frac{A\omega^2 e^{-kt}}{(\omega^2 + k^2)^2}\{k^2(1 + kt) - \omega^2(1 - kt)\}$$

where

$$\tan \alpha = \frac{2\omega k}{\omega^2 - k^2}.$$

We shall now consider three cases with the undamped seismograph (a) when the period of the seismograph is greater than that of the waves in the ground (b) when the period of the seismograph is equal to that of the waves and (c) when the period of the seismograph is less than that of the waves.

In case (a) let $\omega=2\pi/12$, $\mu=2\pi/36$ and $A=1$ so that

$$y = -(8/9)(\cos (30t)^\circ - \cos (10t)^\circ)$$
$$x = 1 - \cos (30t)^\circ.$$

Fig. 3 shows the curves given by these equations. We see that the movement of the ground sets up an oscillation of the seismograph so that the deflections

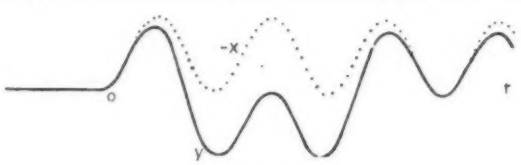

Fig. 3.

obtained are the resultants of the oscillations of the ground and of the seismograph.

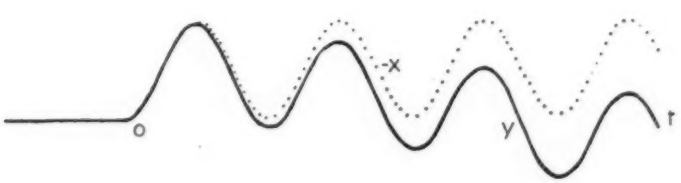

Fig. 4.

Fig. 4 shows the curves obtained with $\omega=2\pi/12$ and $\mu=2\pi/120$ so that $y=-\cos(3t)^\circ+\cos(30t)^\circ$ $x=1-\cos(30t)^\circ$. When, as in this case, the period of the seismograph is ten times that of the waves, the undamped seismograph records the first wave fairly accurately.

Now consider case (b) when the period of the seismograph is equal to that of the waves. If we put $\omega=\mu+\epsilon$ in the equation

$$y = \frac{A\omega^2}{\omega^2 - \mu^2}(\cos \omega t - \cos \mu t)$$

we get when ϵ is very small $y = -A\mu t/2 \sin \mu t$. If $A = 1$ and $\omega = \mu = 2\pi/12$ this gives

$$-y = \frac{\pi}{12} t \sin (30t)^\circ$$

and

$$x = 1 - \cos (30t)^\circ.$$

Fig. 5 shows the curves for y in this case. The amplitude of the first oscillation is only about one half that of the ground but the amplitude rapidly increases as the successive waves arrive.

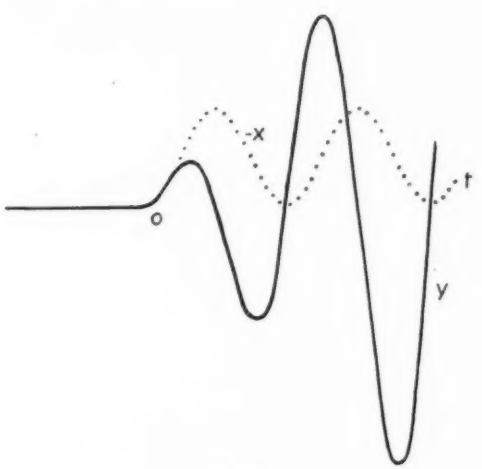

Fig. 5.

We see from this that for the purpose of detecting the arrival of the beginning of a train of waves, it is not a good plan to use a seismograph having a free period equal to that of the waves. A seismograph with a longer period gives twice as great a deflection due to the first wave as one with a period equal to that of the waves. Of course at the start $-y = x$ in all cases so that the initial sensibility is independent of the periods but it is advantageous to get a larger deflection due to the first wave since a large deflection is less likely to be overlooked or obliterated by the small oscillations due to local disturbances.

Now consider the third case when the period of the seismograph is smaller than that of the waves. The equation

$$y = \frac{A\omega^2}{\mu^2 - \omega^2}(\cos \mu t - \cos \omega t)$$

shows that making μ/ω large will diminish the maximum values of y very greatly. For example if $\mu/\omega = 10$ we get

$$y = (A/99)(\cos 10\omega t - \cos \omega t)$$

so that the maximum amplitude is only about one percent of that of the waves in the ground. In this case we get short oscillations due to the term $\cos 10\omega t$ superposed on the curve given by $-(A/99)\cos \omega t$. Hence a smooth curve drawn across the small oscillations would give $-(A/99)\cos \omega t$ which resembles the curve given by $x = A(1-\cos\omega t)$ although displaced downwards and with only one percent of the amplitude. This suggests that the results obtained with high frequency undamped seismographs can be improved by drawing a smooth curve through the small oscillations of frequency equal to that of the seismograph.

Fig. 6 shows the curves given by

$$x = 1 - \cos (30t)^\circ$$

and

$$-y = (\tfrac{1}{8})(\cos (90t)^\circ - \cos (30t)^\circ)$$

which are those for the case $A = 1$, $\omega = 2\pi/12$ and $\mu = 2\pi/4$ in which the period of the seismograph is one third that of the waves.

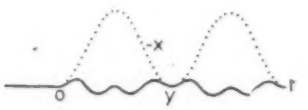

Fig. 6.

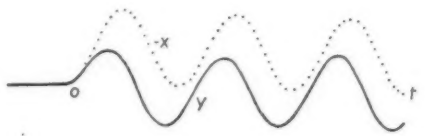

Fig. 7.

Now consider the case of a critically damped seismograph. In this case if $\omega = 2\pi/12$ so that $x = 1-\cos(30t)^\circ$ and $k = p = 2\pi/36$ so that the seismograph period is three times that of the waves then the equation for y gives

$$y = \frac{9}{10}\cos (30t + 36{\cdot}87)^\circ + \frac{5\pi t - 72}{100}e^{-\pi t/18}.$$

Fig. 7 shows the curves given by these equations. Comparing Fig. 7 with Fig. 3 we see that the critical damping cuts out the oscillations of the seismograph, set up at the start, so that we do not get interference between two

Fig. 8.

vibrations of different frequencies. The damping however diminishes the interval at the start in which x and $-y$ agree. Thus without damping in Fig. 3 the two curves agree from $t=0$ to $t=4$ but with critical damping in Fig.

7 they only agree from $t=0$ to $t=1$. We see that to obtain the actual motion of the ground it is necessary to use a much longer period with a critically damped seismograph than with an undamped one.

Fig. 8 shows the curves for x and $-y$ when $\omega=p=k$ so that the periods of the waves and of the seismograph are equal. The equations for x and y in this case are

$$x = 1 - \cos(30t)^\circ$$
$$y = -\tfrac{1}{2}\sin(30t)^\circ + (\pi t/12)e^{-\pi t/6}.$$

The seismograph is less sensitive and x and $-y$ do not agree as well as when the period of the seismograph is greater than that of the waves.

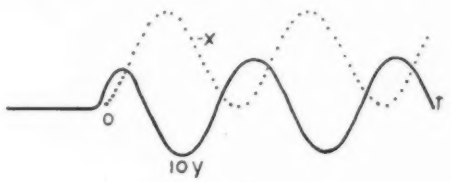

Fig. 9.

If the period of the seismograph is one third that of the waves and the damping critical then we may take

$$x = 1 - \cos(30t)^\circ$$

so that $\omega=\pi/6$ and $k=\pi/2$. The equation for y is then

$$y = -\frac{1}{10}\cos(30t - 36{\cdot}87)^\circ + \frac{5\pi t + 8}{100}e^{-\pi t/2}.$$

Fig. 9 shows the curves for x and $-10y$ in this case. $-10y$ is plotted instead of $-y$ because the maximum value of y is only one twentieth that of the maximum value of x.

Comparing Fig. 9 with Fig. 7 we see that the deflections of a critically damped seismograph with period three times that of the waves are somewhat similar in form but about ten times larger than those of a critically damped seismograph with period one third that of the waves.

It appears that, while in all cases the initial deflections are equal and opposite to the motion of the ground, the only case in which this remains true approximately for more than a fraction of a wave is that in which the period of the seismograph is large compared with that of the waves and the seismograph is undamped.

The ratio of the maximum value of y to the maximum value of x is about $\omega^2/(\omega^2-\mu^2)$ or if T is the period of the waves and T' that of the seismograph it is $T'^2/(T'^2-T^2)$. This is equal very nearly to unity when T'/T is large and is nearly equal to $(T'/T)^2$ when T'/T is small. When T' and T are nearly equal it becomes large but not for the first wave.

Very short period seismographs are difficult to make sufficiently sensitive

because of the small value of the ratio of the maximum value of y to that of x. A critically damped seismograph with a period very small compared with that of the waves in the ground gives a deflection very nearly proportional to the acceleration of the ground. The equation

$$y = \frac{1}{\mu} \int_0^t e^{-k(t-\eta)} \sin \mu(t - \eta) F(\eta) d\eta$$

when $\mu = 0$ for critical damping becomes

$$y = \int_0^t e^{-k(t-\eta)} (t - \eta) F(\eta) d\eta.$$

If the period of the seismograph is very small so that k is large then the factor $e^{-k(t-\eta)}$ is negligible except when η is nearly equal to t. If we suppose that $F(\eta)$ varies slowly and so can be regarded as constant during the short interval when η is nearly equal to t then

$$y = F(t) \int_0^t e^{-k(t-\eta)} (t - \eta) d\eta$$

or

$$y = \frac{F(t)}{k^2} (1 - e^{-kt}(1 + kt))$$

or when t is not very small $y = F(t)/k^2$. Thus in this case since $F(t) = -\ddot{x}$ we have $y = -\ddot{x}/k^2$.

If a very short period critically damped seismograph can be made sufficiently sensitive, it will indicate the acceleration of the ground with considerable accuracy. Whenever a sudden change in the acceleration of the ground occurs, the change in the seismograph deflection will be proportional to the change in the acceleration multiplied by the factor $1 - e^{-kt}(1 + kt)$ where t is the time since the sudden change took place, which very soon becomes equal to unity when k is large.

Now consider the deflection of the seismograph due to a single wave. Let $x = A(1 - \cos \omega t)$ from $t = 0$ to $t = 2\pi/\omega$ and $x = 0$ when $t < 0$ and when $t > 2\pi/\omega$. In this case y is given by

$$y = -\frac{A\omega^2}{\mu} \int_0^t e^{-k(t-\eta)} \sin \mu(t - \eta) \cos \omega\eta d\eta$$

from $t = 0$ to $t = 2\pi/\omega$ and by

$$y = -\frac{A\omega^2}{\mu} \int_0^{2\pi/\omega} e^{-k(t-\eta)} \sin \mu(t - \eta) \cos \omega\eta d\eta$$

when t is greater than $2\pi/\omega$ because the acceleration is zero when $t > 2\pi/\omega$.

First suppose that the seismograph is undamped so that $k = 0$ and then when $t > 2\pi/\omega$ we have

$$y = -\frac{A\omega^2}{\mu} \int_0^{2\pi/\omega} \sin \mu(t - \eta) \cos \omega\eta d\eta$$

which gives

$$y = \frac{2A\omega^2}{\omega^2 - \mu^2} \sin \frac{\mu\pi}{\omega} \sin \mu(t - \pi/\omega).$$

From $t=0$ to $t=2\pi/\omega$ the equation for y is

$$y = \frac{A\omega^2}{\omega^2 - \mu^2}(\cos \omega t - \cos \mu t)$$

as for $x=A(1-\cos \omega t)$ when $t>0$.

The curves for y and x in the three cases

$$A = 1, \quad \omega = \pi/6, \quad \mu = \pi/18$$
$$A = 1, \quad \omega = \pi/6, \quad \mu = \pi/6$$
$$A = 1, \quad \omega = \pi/6, \quad \mu = \pi/2$$

are shown in Fig. 10, Fig. 11, and Fig. 12.

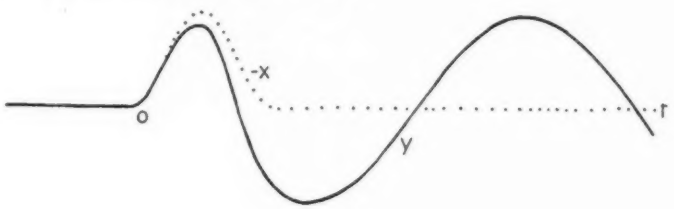

Fig. 10.

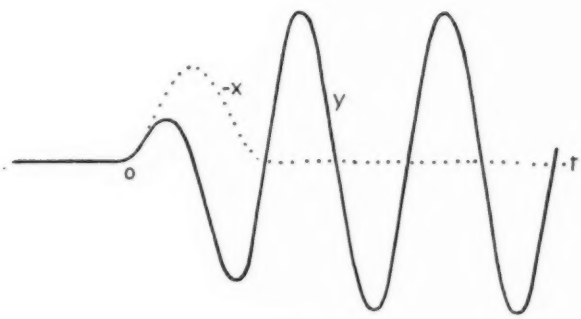

Fig. 11.

Fig. 12.

In Fig. 10 the period of the seismograph is three times that of the wave and a large oscillation is started in the seismograph. In Fig. 11 the two periods are equal. In Fig. 12 the period of the seismograph is one third that of the wave and the wave ends with the seismograph in equilibrium so that $y=0$ after the wave has passed.

Now suppose that the seismograph is critically damped so that $\mu=0$ and $k=p$. In this case the equation for y when $t>2\pi/\omega$ is of the form $y=a+b(t-2\pi/\omega)e^{-k(t-2\pi/\omega)}$ and the curves from $t=0$ to $t=2\pi/\omega$, of course, are the same as with $x=A(1-\cos \omega t)$ when $t>0$. Figs. 13, 14 and 15 show the curves for $-y$ and x in the three cases

$$A = 1, \quad \omega = \pi/6, \quad \mu = \pi/18$$
$$A = 1, \quad \omega = \pi/6, \quad \mu = \pi/6$$
$$A = 1, \quad \omega = \pi/6, \quad \mu = \pi/2.$$

We see that with critically damped seismographs the shape of the curves obtained depends less on the period of the seismograph than with undamped seismographs. When the period of the seismograph is much smaller than that of the wave the deflections are very small.

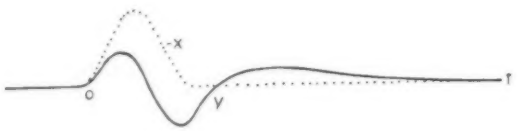

Fig. 13.

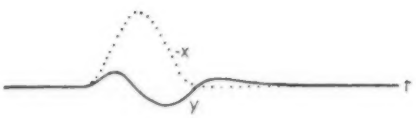

Fig. 14.

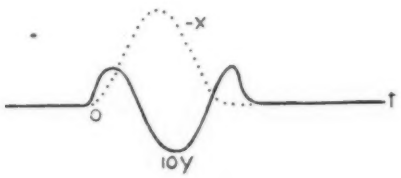

Fig. 15.

It appears that the only case in which the deflection of the seismograph is approximately equal and opposite to that of the ground during the passage of one or more waves is that of an undamped seismograph with period much larger than that of the waves. In this case, however, a single wave sets up an oscillation which continues indefinitely.

The principal conclusion which can be drawn from the above results is that seismographs do not register the actual motion of the ground so that arguments based on wave forms or the apparent arrival of waves later than the first wave must be used with great caution.

Critically damped seismographs are obviously better when it is desired to observe the arrival of several waves separated by appreciable intervals.

The inverse problem of calculating the motion of the ground from the recorded deflections of the seismograph will now be considered. The equation $-\ddot{x}=\ddot{y}+2k\dot{y}+p^2y$ on integrating gives $-\dot{x}=\dot{y}+2ky+p^2\int_0^t ydt+C$. We may suppose that $y=0$ at $t=0$ so that $-\dot{x}_0=\dot{y}_0+C$. But at $t=0$ $-\dot{x}_0=\dot{y}_0$ so that $C=0$. Also when t is very large we may suppose $\dot{x}=\dot{y}=y=0$ so that $\int_0^\infty ydt=0$. Integrating again we get

$$-x = y + 2k\int_0^t ydt + p^2\int_0^t \left\{\int_0^t ydt\right\}dt + D.$$

If $x=y=0$ at $t=0$ then $D=0$. Let $\int_0^t ydt=a_1$, and $\int_0^t\{\int_0^t ydt\}dt=a_2$ so that $-x=y+2ka_1+p^2a_2$. When t is large we may suppose that $x=y=0$ so that we must have $2ka_1+p^2a_2=0$. But $a_1=0$ when $t=\infty$ so that $a_2=0$ when $t=\infty$.

In the equation $-x=y+2ka_1+p^2a_2$ since $p^2=4\pi^2/T^2$ where T is the natural period of the seismograph, when undamped, we see that the correction p^2a_2 which must be added to y to get the true motion of the ground is inversely as T^2 for all values of the time t. If T is very large this correction may be negligible. In the same way the damping correction $2ka_1$ is proportional to the damping coefficient. Thus for an undamped seismograph of very long period $-x=y$ so that the seismograph gives the ground motion exactly. Any increase of the damping and any reduction of the period T increases the corrections and so increases the difference between the seismograph deflections and the ground motion. If $k=p$ the seismograph is critically damped.

To find a_1 it is convenient to make a tracing of the seismogram on graph paper and draw a curve for the area by counting the squares between the y curve and the zero line $y=0$. The zero line is usually not marked on the seismograms so that it is necessary to draw it. A small error in the position of the zero line makes a large error in a_1. Thus if it is drawn at $y=-\frac{1}{2}$ instead of at $y=0$ we get a_1 too large by $\frac{1}{2}t$. Such errors usually occur but are easily allowed for when it is remembered that $a_1=0$ for large values of t. If a_1 does not finally approach zero as t increases then by slightly shifting the zero line it is made to do so. Very often it is found that the area curve fluctuates about a straight line through the origin finally approaching this line. In such cases it is clear that the zero line is too high or too low. The line through the origin may then be taken to be the zero line for the area. The integral a_2 is then obtained from the curve giving the area in the same way. If it does not finally approach zero it must be made to do so by a suitable shift of its zero line. If an error Δa is made in the area a_1 at time t_1 then this produces an error $\Delta a(t-t_1)$ in the value of a_2 at a time t later than t_1. Such errors are usually appreciable and cause the curve obtained for this integral to drift up and down. It is therefore usually necessary to draw a new zero line for this curve. This zero line must be a smooth curve such that the value of the integral oscillates more or less equally above and below it.

Instead of endeavouring to correct such errors by shifting the zero lines, a process which depends a good deal on arbitrary judgment, one may carry out the calculation of x using the uncorrected curves for the area and the integral. When this is done the curve obtained for the motion of the ground usually appears to oscillate about a smooth curve which may drift very far from the zero line. This smooth curve may be drawn and taken to be the true zero line for the ground motion. These methods of correcting the calculated ground motion are necessary unless the areas are very accurately determined. They obviously result in the removal from the ground motion of any long period oscillations having periods long compared with the period of the seismograph. The constant p is nearly equal to $2\pi/T$ unless the damping is large, T being the periodic time of the seismograph. In getting the areas of the curves it is convenient to use a time scale such that $p=1$.

Before considering results obtained with seismograms some simple theoretical cases will be considered. Suppose $y=0$ when $t<0$ and $y=a \sin pt$ when $t>0$ and that $k=0$. In this case

$$-x = a \sin pt + ap^2 \int_0^t \left\{ \int_0^t \sin pt\, dt \right\} dt$$

which gives

$$-x = a \sin pt + ap \int_0^t (1 - \cos pt) dt$$

so that

$$-x = a \sin pt + pat - a \sin pt = pat.$$

Thus in this case the ground moves with uniform velocity ap. The sudden starting of this motion at $t=0$ starts the seismograph oscillating freely with amplitude a but has no further effect on the motion. If a ground wave ar-

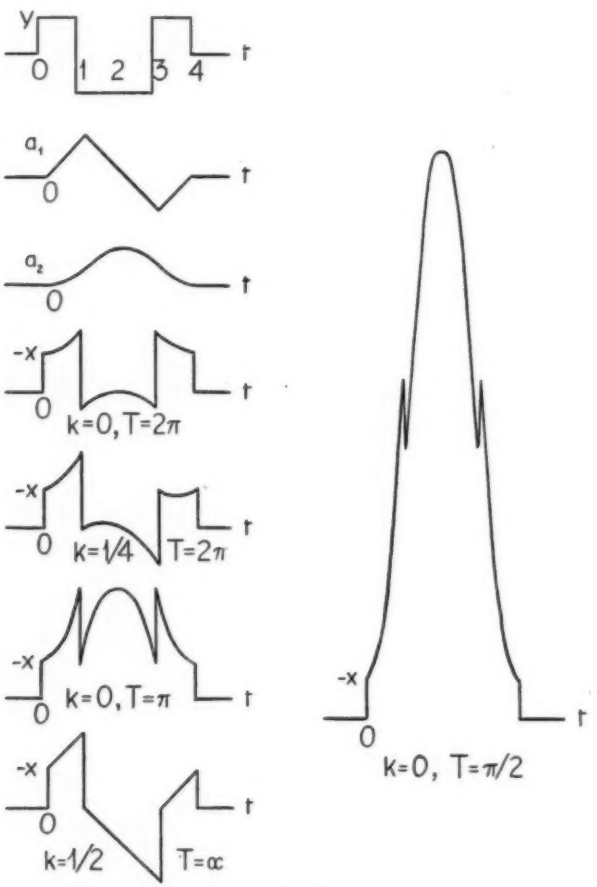

Fig. 16.

rives the period of which is long compared with that of the seismograph it will produce a nearly uniform motion of the ground for a short time which will start the seismograph oscillating with its natural period.

Suppose that $y=0$ when $t<0$ and $y=a\cos pt$ when $t>0$ and $k=0$. In this case $-x=a\cos pt+ap^2\int_0^t\{\int_0^t \cos ptdt\}dt$ which gives $-x=a$. Thus in this case the ground suddenly moves a distance a at $t=0$ and then remains at rest.

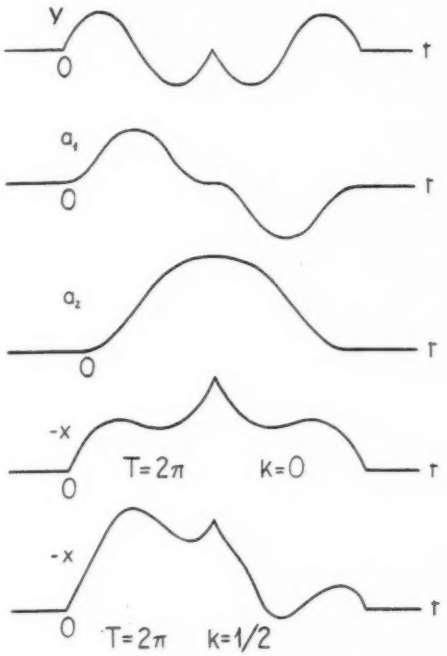

Fig. 17.

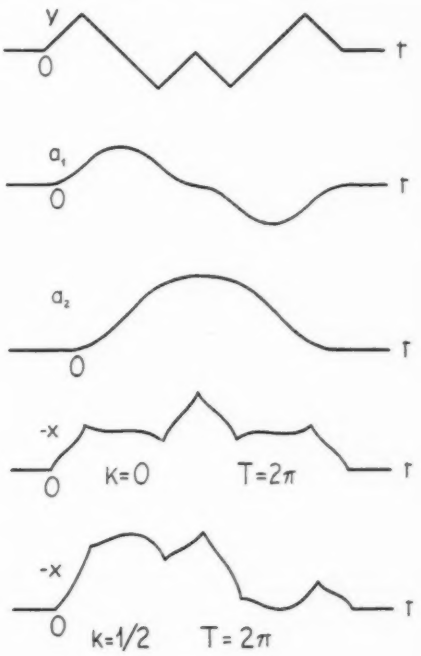

Fig. 18.

Figs. 16, 17, 18, show the curves for y, a_1, a_2, and $-x$ in some simple theoretical cases which need no explanation. The seismograph of course always records sudden movements correctly. Fig. 19 shows a seismogram of the waves due to an explosion of 1000 pounds of dynamite at a distance of 9870 meters. The seismograph was of the vertical mechanical type. The calculated

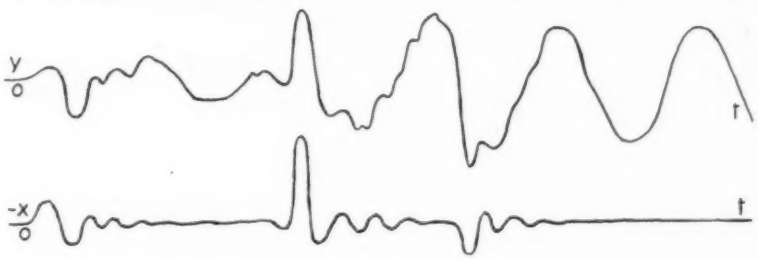

Fig. 19.

curve for the ground motion (x) is also shown. It appears that the explosion waves consist of a rapidly damped oscillation of period about 0.1 sec. Several such waves come in at different times and all have about the same shape and period. In this case there were no geological anomalies in the neighborhood, but the wave velocity increased more or less uniformily with the depth. The

waves therefore travelled along nearly circular paths and may have been reflected from the ground surface as shown in Fig. 20. In this way the arrival of several waves at different times may be explained. Fig. 21 shows a seismogram of the waves from 1000 pounds of dynamite at 9590 meters. In this case

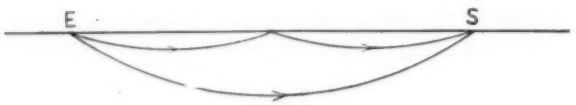

Fig. 20.

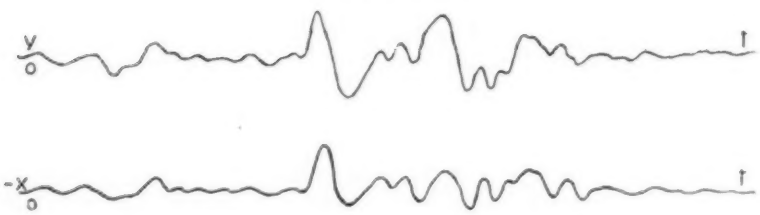

Fig. 21.

there was a salt dome in the path of the waves. The curve for x the ground motion is also shown. The ground motion can be regarded as mainly due to several rapidly damped oscillations which, however, are not as well separated as in the previous case.

An integraph[2] has been constructed which enables the curves for a_1 and a_2 to be drawn quickly. Fig. 22 is a diagram of this integraph. FF' is a brass block which rolls on two wheels WW' and carries a pencil at its middle point x. The points FF' on this block are connected to two parallel bars AD and $A'D'$ by means of two Peaucellier linkages which allow F and F' to move along straight lines perpendicular to AD and $A'D'$ respectively. The bars

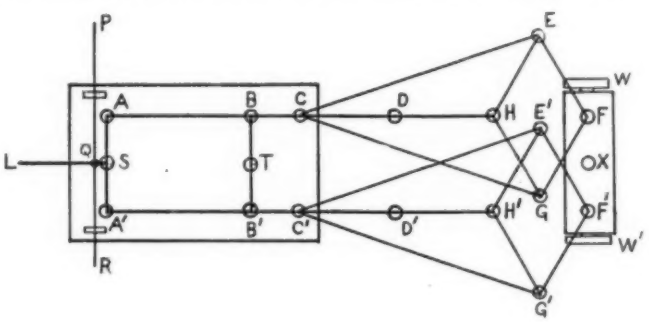

Fig. 22.

AD and $A'D'$ are connected by two links AA' and BB' which are supported on pivots at S and T fixed to a base. The link AA' can be turned about S by means of a liver SL. This lever also moves a bar RQP which can slide on the

[2] Intergraphs based on the same general principle are described in Les Intergraphs. Abdank-Abakauowiz. Paris, (1886).

base parallel to itself. When AA' is turned the block FF' also turns so that it is always parallel to AA'. When the lever SL is turned through an angle θ the end P of the bar PR moves a distance $l \tan \theta$ where l is the distance from S to PR.

The integraph is moved along parallel to the zero line of the curve of which the area is desired and the point P is made to follow the curve by moving the lever SL. When the base is moved a distance dx the pencil moves along a distance ds such that $ds \cos \theta = dx$. Let $ds \sin \theta = dz$ so that $dz/dx = \tan \theta$. If the first curve is given by $y = f(x)$ then $y = l \tan \theta$. Hence $y = ldz/dx$ or $ydx = ldz$ so that $lz = \int ydx \cdot z$ is the distance the pencil moves side ways so that the pencil draws a curve giving the area of the first curve.

In the instrument which was made the lengths AA', BB', FF' were each two inches and l was made one cm. The links were made of stainless steel 3/8 by 1/16 inch. The wheels were 7/8 inch ball bearings. The block FF' was $3 \times 1 \times 1$ inches.

This instrument enables the curves for a_1 and a_2 to be drawn quickly and with sufficient accuracy for most purposes. In using it one side of the base is kept against a straight edge fastened to the drawing board.

EDITOR'S COLUMN

The Velocity of Sound in Gases

IN THE classical theory of the propogation of sound in gases it is assumed that the gas acts like an elastic medium. The general theory of the propagation of pressure disturbances (of which sound is one form) through such a medium shows that the square of the velocity of propagation is equal to the rate of change of the pressure of the medium with the density (i.e., $dp/d\rho$). The formula connecting the pressure with the density (or volume) of a gas depends upon the supplementary conditions specifying the state of the gas; e.g. whether the process takes place at constant temperature (isothermally) or without the transfer of heat (adiabatically). It depends also upon the equation of state of the gas, of course, but this dependance is not so important at temperatures not too near the liquefaction point of the gas. By comparison with the experimental values for the velocity of sound and also by a study of the kinetic theory of gases, one concludes that in the rapid pulsations occurring in the gas during the propagation of sound waves of ordinary frequencies (40∽10,000 cycles/sec.) the gas behaves adiabatically as the time interval between oscillations is too short for a transfer of heat between the various portions of the gas. This leads to the formula

$$v = [\gamma p/\rho]^{1/2}$$

where v is the velocity of sound, p the pressure, ρ the density, and γ is the ratio of the specific heat at constant pressure to that at constant volume. For simple monatomic gases such as He the value of γ is about 1.67 while for more complicated gases it lies between 1.67 and 1, being closer to the latter value the more complex the structure of the molecule, generally speaking. In fact, the law of equipartition of energy of Maxwell shows that

$$\gamma = 1 + 2/s$$

where s is the number of independent types of motion which the molecule exhibits. While this last formula is no longer supposed to be altogether correct, if properly interpreted it does give one a clue as to the modes of vibration of molecule. If one then measures the velocity of sound, and supposes its formula as given above to be sufficiently substantiated theoretically, he can use his measurements for a calculation of γ, and so can gain some insight into the complexity of the molecule.

Recent measurements of this nature have been made by H. Kneser[1] and discussed from the point of view of molecular dynamics by O. Heil.[2] Kneser found that while the velocity of sound at ordinary frequencies in carbon dioxide gas is about 258.5 meters/second it increases to about 268.2 meters/second for sound frequencies around 10^5/sec. This implies a change in γ from about 1.3 to 1.4, and a consequent change in s of roughly one unit. Presumably then, for very high frequencies some mode of vibration of the molecule is not excited.

According to the chemical facts concerning CO_2, the molecule seems to have a linear structure, i.e., the equilibrium position of the two O-atoms is on a straight line through the C-atom and equidistant from it. The easiest mode of vibration to excite in this system, and the only one which will be appreciably excited at room temperatures, is the oscillation of the C-atom normal to the line joining the O-atoms, the energy of oscillation of the O-atoms themselves being much larger and consequently harder to excite. It is known, however, that it is sometimes quite difficult to transfer energy of translational motion into energy of vibrational motion in the collision between two molecules because of the necessity for the conservation of momentum as well as of energy. In the particular case of CO_2 it appears that only one collision in 10^6 is effective in exciting this vibration of the C-atom. Because of this there is not a sufficient number of collisions occurring during the time occupied by a sound vibration to excite this oscillation, and as a consequence the gas acts effectively as though its molecules were rigid bodies with one moment of inertia (that about the common line of the atoms) zero, so that

[1] H. Kneser, Phys. Zeits. **32,** 179 (1931).
[2] O. Heil, Zeits. f. Physik **74,** 31 (1932).

the effective number of degrees of freedom of the system is 5; i.e., $\gamma = 1+2/5 = 1.4$. In measurements of the specific heat the time required for making such a measurement is naturally so long that equilibrium conditions can be established within the gas, and this effect does not occur. This intrusion of molecular phenomena into a large scale phenomenon like the propagation of sound in a gas furnishes an interesting example of the direction of motion of the science of physics.

E. L. HILL

The Pauli Exclusion Principle

EVER since the enunciation of the famous "exclusion" principle by W. Pauli in 1925 there has been so much discussion about it,[1] and it has been so fully confirmed from such a variety of data that it has become one of the most firmly established of physical principles. While we cannot hope to bring out and establish detailed applications of the principle in the following lines, we shall attempt to make its meaning a bit clearer to those to whom a purely mathematical statement would mean little.

It is important to realize that Pauli arrived at the statement of his principle from a study of experimental data before the development of the quantum mechanics—from the data on the character of the orbits occupied by the electrons in complicated atoms. In the atomic theory developed by Bohr the attempt was made to assign each electron to a certain definite orbit which could be specified by quantum numbers. To each orbit there was assigned a set of three quantum numbers n, k, and m. From a consideration of the actual way in which the data on the spectra of the atoms showed these orbits to be occupied, Pauli concluded that not more than two electrons would ever be found to occupy such an orbit. The picture of an electron orbit according to this theory was an ellipse around which the electron travelled rapidly while the nucleus was stationed at one focus of the ellipse. The integer n determined the major axis of the ellipse, k the minor axis, while m determined the orientation of the normal to the plane of the ellipse in space.

In order to distinguish between the two electrons which could occupy such an orbit or "state," the next step was to introduce a fourth quantum number. Thus the state of an individual electron is now specified by four quantum numbers n, l, m_l, m_s; the first three have nearly the same significance as in Bohr's theory but the fourth m_s, is the so-called spin "magnetic" quantum number which assumes the two values ($\pm\frac{1}{2}$) just sufficient to distinguish the two orbits having the same values of n, l, and m_l.

It is sometimes said that each *electron* in an atom is specified by certain quantum numbers. This way of speaking is based on the older theory in which each individual electron was supposed to occupy a given orbit in the strict mechanical sense of the word, but it is now anachronistic and should not be used when there is any danger of being misunderstood. Quantum mechanics as it is understood at present does not give one the slightest information about what the individual electrons may be up to, but instead it at most gives one information concerning the possible states which the electrons can occupy. It is the determination of these *allowed states* which lie at the basis of the theory. These states are determined, we suppose, from a "wave equation" of some form, either as specified by Schrödinger, or perhaps in the (still unsettled) relativistic form of Dirac.

It is a general rule in quantum mechanics (although this point has scarcely been studied at all from the mathematical point of view) that there are as many quantum numbers for the specification of a state of a system as there are degrees of freedom (i.e. independent coordinates) in the system. Since it is found that one requires four quantum numbers for the complete specification of the individual states occupied by an electron in an atom, one would suppose that an electron has four degrees of freedom, three of which are certainly the three degrees of freedom corresponding to the Cartesian coordinates of the center. The fourth degree of freedom is now interpreted as due to the fact that the electron does not act altogether like a point charge but has a spinning motion like a top. In accordance with the general principles of quantum mechanics one coordinate is required for the specification of the direction of the axis of the top in space, that is, the space "quantization" of the spin

[1] W. Pauli, Zeits. f. Physik **31,** 765 (1925).

angular momentum vector. This is the rôle of the quantum number m_s which measures the component of the spin angular momentum (in units $h/2\pi$) along some preassigned axis in space. In reality a fifth quantum number is necessary for the specification of the magnitude of the spinning motion of the electron, but it seems that ordinarily all electrons have the same amount of spinning motion around the axis of spin, and consequently this fifth degree of freedom does not play any important rôle. In the relativistic theory of the electron, however, the situation is not so simple.

On the basis of this rather general theory it is to be expected, then, that four quantum numbers will be needed for the specification of the state of an electron even when that electron does not form part of an atom; i.e., even when it is part of a molecule or of a large crystal. This raises the question: will there be any analogue of Pauli's principle applicable to electrons which are not part of an atom? Clearly there must be, for it can hardly be allowed that an atom represents any particularly peculiar kind of mechanical system. What then, is this analogue or generalization of the original exclusion principle?

The answer which theoretical physicists have given to this question may be found by looking for the essential features of the principle as applied to electrons within atoms. This generalization is necessarily of a mathematical nature in order to be of use in a mathematical theory but that does not necessarily mean that it is really complex. When we think of the *states* within an atom as being specified in some way, through the intermediary use of the wave equation, and then of the electrons as being assigned to these states, we see that Pauli's principle states in effect: not more than one electron is to be allowed to occupy a given state (the state being completely specified say by the four quantum numbers n, l, m_l, m_s). This is a general rule which could in principle be applied to *any* problem in which the states which electrons can occupy are completely specified. As an example we may consider the problem of the motion of a set of electrons within a cubical box which has "reflecting" walls preventing the escape of the electrons. Following the original ideas of de Broglie, it is convenient to imagine a set of waves filling the enclosure. The possible "states" for such a system are given by the condition that the associated waves shall form a standing wave system. The wave-lengths of the waves in each of the three directions normal to the sides of the box must accordingly be integral sub-multiples of twice the length of the box ($\lambda = 2a/n$). Three of the quantum numbers defining the *states* can now be taken as the integers n_1, n_2, n_3 specifying the wave-lengths in these three directions. In order to define the state completely we need one more quantum number which is to be taken as the projection of the spin vector of the electron along some arbitrary direction, the quantum number m_s. This gives us just the model of a metal which has been used so successfully by Sommerfeld.[2] The energy of an electron occupying a state specified by the values of the quantum numbers $n_1 = n_1^1$; $n_2 = n_2^1$; $n_3 = n_3^1$; $m_s = m_s^1$ is found to be

$$(h^2/8ma^2)\left[(n_1^1)^2 + (n_2^1)^2 + (n_3^1)^2\right]$$

and to this approximation does not depend upon m_s^1.

The generalized Pauli principle would then tell us that not more than one electron can occupy a state specified by the four quantum numbers n_1, n_2, n_3, m_s. This restriction would be expected to apply regardless of the magnitude of a, that is, regardless of the size of the crystal or piece of metal, and is thus not specifically a result depending upon the smallness of atoms and molecules. A crystal or a piece of metal (or even a star) is just as truly a single quantum mechanical system as is a helium atom.

While this method of stating Pauli's principle in terms of states defined by means of quantum theory is general, it is by no means always convenient, and has been replaced by a more succinct and clear statement to the effect that the complete wave function specifying the state of the system (where "the system" now means the whole group of electrons and protons), must be *antisymmetrical* in the coordinates of each pair of electrons—if the coordinates of any two electrons be interchanged in the wave function the latter must be changed in sign. In the "coordinates" of an electron must be counted a spin variable analogous to m_s as well as the three Cartesian coordinates of the center of the electron. The particular advantage of this specification is that one

[2] E. g., A Sommerfeld and N. Frank, Rev. Mod. Phys. **3**, 1 (1931).

need not bother (unless he finds it convenient) to specify the actual states which the individual electrons are to occupy.

There is no further theoretical justification of the exclusion principle than this. It is not a "deriveable" principle—but like wave mechanics itself, draws its justification from its agreement with the experimental facts. That it has been found to fit in so nicely with wave mechanics is, of course, no proof of its validity or explanation of its physical content, but nevertheless does give one a more secure feeling concerning it.

It has been found possible to extend the scope of the principle to include the general case in which the system contains various types of similar particles, such as photons, nuclei, etc. For example, it is found that photons apparently obey an "inclusion" principle in that great numbers of them tend to crowd into the same state.

E. L. Hill